



HAL
open science

Reconstruction 3D de bâtiments en interférométrie RSO haute résolution: approche par gestion d'hypothèses

François Cellier

► **To cite this version:**

François Cellier. Reconstruction 3D de bâtiments en interférométrie RSO haute résolution: approche par gestion d'hypothèses. domain_other. Télécom ParisTech, 2007. English. NNT: . pastel-00002228

HAL Id: pastel-00002228

<https://pastel.hal.science/pastel-00002228>

Submitted on 9 Mar 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ecole nationale supérieure des télécommunications

THESE

pour obtenir le grade de

Docteur de l'école nationale supérieure des télécommunications

Spécialité : "Signal et Images"

préparée à l'Office Nationale d'Etudes et de Recherches Aérospatiales
au sein du Département de Traitement de l'Information et Modélisation
et du Département ElectroMagnétisme et Radar

présentée et soutenue publiquement

par

François Cellier

le 10 janvier 2007

Titre :

Reconstruction 3D de bâtiments en interférométrie RSO haute résolution : approche par gestion d'hypothèses

Directeur de thèse : Jean-Marie Nicolas

Encadrant ONERA : Hélène Oriot

Jury :

Monsieur U. Stilla	Professeur du TUM (Allemagne)	Président
Monsieur X. Descombes	Chargé de recherche à l'INRIA	Rapporteur
Monsieur E. Pottier	Professeur de l'université Rennes 1	Rapporteur
Monsieur M. Datcu	Professeur de l'ENST Paris	Examineur
Madame H. Oriot	Ingénieur de Recherche à l'ONERA	Examineur
Monsieur J.M. Nicolas	Professeur de l'ENST Paris	Directeur de thèse

Remerciements

Cette thèse s'est effectuée à l'ONERA au sein du département Traitement de l'Information et Modélisation (Chatillon) et du département ElectroMagnétisme et Radar (Palaiseau). Je remercie donc Messieurs Gilles Foulon et Jean-Marc Boutry de m'avoir accueilli au sein de leur département respectif et de m'avoir permis d'effectuer cette thèse dans les meilleures conditions possibles.

Je remercie tout particulièrement Madame Hélène Oriot, ingénieur de recherche à l'ONERA, et Monsieur Jean-Marie Nicolas, professeur à l'ENST, de m'avoir encadré durant cette thèse. Je remercie tout d'abord Hélène de m'avoir choisi pour cette thèse et permis de travailler dans le domaine passionnant de l'imagerie RSO. Je la remercie également pour sa gentillesse, pour sa patience et pour tous ses conseils avisés dans les moments importants. Je remercie également Jean-Marie pour sa gentillesse, pour avoir suivi avec intérêt mes travaux tout au long de cette thèse, notamment lors de ces réunions matinales à l'ENST. Enfin, je remercie Hélène et Jean-Marie pour leur soutien permanent et pour m'avoir appris à valoriser mes travaux.

Je voudrais ensuite remercier les membres du jury qui ont évalué ces travaux de thèse. Tout d'abord je remercie Messieurs Xavier Descombes, chargé de recherche à l'INRIA, et Eric Pottier, professeur de l'université Rennes 1, d'avoir pris le temps de rapporter ce manuscrit et pour leurs remarques judicieuses sur mes travaux. Je voudrais également remercier les professeurs Uwe Stilla (TUM) et Mihai Datcu (DLR / ENST) d'être venus d'Allemagne pour évaluer mon travail.

Au cours de cette thèse, j'ai été amené à travailler avec diverses personnes. Tout d'abord, je voudrais remercier Elise Colin Koenieger qui m'a fait découvrir la polarimétrie interférométrique RSO. Je la remercie pour ses qualités pédagogiques et son honnêteté scientifique. Je remercie également Anne Beaupère pour son aide sur le géo-référencement et Patrick Secchi pour son aide sur divers problèmes informatiques.

Je tiens également à remercier les personnes qui m'ont permis d'acquérir une vision plus générale du radar. Tout d'abord, je remercie Philippe « Michel » Martineau et Hubert Cantalloube d'avoir pris le temps de répondre à mes nombreuses questions. Je remercie aussi Florence Tupin de l'ENST et Frédéric Galland de l'institut Fresnel pour l'organisation des séminaires radar ainsi que l'ensemble des participants à ces séminaires. Grâce à eux j'ai pu découvrir de nombreux domaines de recherche en imagerie RSO.

Je voulais également remercier tout particulièrement Xavier Dupuis, Pascale Dubois-Fernandez et Colette Coulombeix de l'unité RIM de m'avoir fourni les images RSO et les vérités terrains associées.

Je remercie tous les membres du DTIM, du DEMR et de l'ENST qui se sont intéressés à mes travaux de thèse pour leur aide et leurs conseils.

Un grand merci au duo Fabrice Savignol / Fabrice Janez ainsi qu'à Françoise Pradines pour leur soutien, leur gentillesse et leur bonne humeur permanente.

Je remercie l'ensemble des docteurs, thésards, apprentis et stagiaires de l'ONERA avec qui j'ai partagé de très bons moments. Je pense tout particulièrement à Vivien, Jojo, Benjamin, Benji, Cédric, Martial et Gilles. Je voulais également remercier les thésards de l'ENST que j'ai vu moins souvent mais qui ont toujours été très accueillants : Antonio, Julien, Pénélope, Cédric et Vincent. Merci également aux rockeurs de l'ONERA, à savoir Ben, Mag, Luc, Xav et JL, d'avoir mis le feu lors de nos répétitions et concerts.

Enfin un grand merci à ma famille et mes amis de s'être intéressés à mes travaux et pour leurs encouragements permanents.

INTRODUCTION GENERALE.....	9
PARTIE 1 : RECONSTRUCTION DE BATIMENT EN IMAGERIE RADAR A SYNTHESE D'OUVERTURE INTERFEROMETRIQUE HAUTE RESOLUTION	13
1. L'IMAGERIE RADAR A SYNTHESE D'OUVERTURE	15
1.1. Introduction	15
1.2. L'imagerie RSO.....	15
1.2.1. Les ondes RADAR.....	15
1.2.2. RADAR et image complexe	16
1.2.3. Définition des résolutions en distance et en azimut.....	17
1.2.4. Les méthodes d'amélioration de la résolution.....	19
1.2.5. Résolution et pas d'échantillonnage.....	24
1.2.6. Conclusion	24
1.3. L'interférométrie RSO	25
1.3.1. Principe.....	25
1.3.2. Phase interférométrique et hauteur	26
1.3.3. Cohérence théorique.....	28
1.3.4. Interférogramme multivue et cohérence empirique	28
1.3.5. Bruits sur la mesure interférométrique.....	28
1.3.6. Conclusion	30
1.4. Conclusion.....	30
2. ETUDE ET INTERPRETATION DES IMAGES RSO HAUTE RESOLUTION EN MILIEU URBAIN.....	31
2.1. Introduction	31
2.2. Géométrie en imagerie RSO.....	31
2.2.1. Géométrie RADAR et géométrie terrain.....	31
2.2.2. Hypothèse d'un front d'onde plan // front d'onde sphérique	31
2.2.3. Distorsions géométriques	36
2.2.4. Les ombres.....	37
2.2.5. Conclusion	37
2.3. Mesures du 3D en imagerie RSO interférométrique	37
2.3.1. Les mesures monoscopiques.....	37
2.3.2. Comparaisons des précisions des mesures monoscopiques et interférométriques ..	40
2.3.3. Conclusion	42
2.4. Analyse phénoménologique	42
2.4.1. Introduction.....	42
2.4.2. L'équation RADAR	42
2.4.3. Notion de rugosité et de textures	43
2.4.4. Les échos	44
2.4.5. Conclusion	63
2.5. Etude des mélanges d'informations 3D au sein d'une zone de repliement.....	63
2.5.1. Origine du repliement.....	63
2.5.2. Modélisation d'un mélange de deux contributions d'altitudes différentes	63
2.5.3. Interprétation du modèle	66
2.5.4. Validation du modèle	68

2.5.5. Limites du modèle.....	71
2.5.6. Conclusion.....	71
2.6. Conclusion.....	71
3. ETUDE BIBLIOGRAPHIQUE SUR LA REPRESENTATION 3D DES BATIMENTS EN IMAGERIE RSO.....	73
3.1. Introduction	73
3.2. Présentation des données.....	74
3.3. Prétraitement.....	75
3.4. Approches surfaciques	76
3.4.1. Extraction du modèle numérique de surface	76
3.4.2. Extraction du modèle numérique de terrain	77
3.4.3. Extraction des zones élevées	78
3.4.4. Conclusion	79
3.5. Approche objet : extraction du modèle numérique d'élévation	80
3.5.1. Introduction.....	80
3.5.2. Modélisation des bâtiments	80
3.5.3. Extraction des primitives.....	80
3.5.4. Extraction de hauteurs	82
3.5.5. Reconstruction	84
3.5.6. Conclusion.....	87
3.6. Conclusion.....	87
PARTIE 2 : PROPOSITION D'UNE METHODE DE RECONSTRUCTION DE BATIMENTS SUR IMAGES RSO INTERFEROMETRIQUE HAUTE RESOLUTION	89
4. EXTRACTION DE SIGNATURES DE BATIMENTS EN IMAGERIE RSO HAUTE RESOLUTION	95
4.1. Introduction	95
4.2. Introduction du filtre mean shift en imagerie RSO	95
4.2.1. Nécessité d'un filtrage des images RSO en amplitude	95
4.2.2. Filtres paramétriques usuels en imagerie RSO.....	96
4.2.3. Le filtre Mean Shift.....	98
4.2.4. Comparaison du filtre mean shift aux filtres paramétriques en imagerie RSO.....	107
4.2.5. Performances sur images synthétiques.....	112
4.2.6. Performances sur images RAMSES	119
4.2.7. Conclusion.....	122
4.3. Utilisation de la cohérence pour le filtrage des ombres	123
4.3.1. Intérêt de la cohérence.....	123
4.3.2. Extension du domaine joint à l'image de la cohérence.....	123
4.3.3. Fusion des images d'amplitude et de cohérence filtrées	124
4.3.4. Performance sur images RAMSES	125
4.3.5. Conclusion	128
4.4. Extraction des signatures de bâtiments en imagerie RSO.....	128
4.4.1. Introduction.....	128
4.4.2. Détection des primitives	128
4.4.3. Description des algorithmes d'extraction.....	129
4.4.4. Résultats	133

4.4.5. Conclusion.....	140
4.5. Conclusion.....	141
5. RECONSTRUCTION DE BATIMENTS PAR GESTION D’HYPOTHESES EN IMAGERIE RSO HAUTE RESOLUTION.....	143
5.1. Introduction	143
5.2. La gestion d’hypothèses pour la reconstruction de bâtiments en imagerie RSO	143
5.2.1. Description hiérarchique d’un bâtiment pour la reconstruction en imagerie RSO	143
5.2.2. Proposition d’une sémantique pour l’interprétation des signatures.....	145
5.2.3. Représentation d’une partie de bâtiment dans cette sémantique	145
5.2.4. Découpage d’un bâtiments en parties élémentaires	146
5.3. Création de l’arbre d’hypothèses	147
5.3.1. Définition de l’arbre d’hypothèse initial	147
5.3.2. Détection de façades avants et arrières	148
5.4. Génération d’hypothèses de bâtiments	149
5.4.1. Génération d’hypothèses sur les ombres.....	149
5.4.2. Génération d’hypothèses entre deux autres segments.....	153
5.5. Validation des hypothèses de bâtiments	154
5.5.1. Gestion des incompatibilités géométriques	154
Dans le cas des ombres, il ne doit y avoir aucune primitive située entre le début et la fin de l’ombre.	155
5.5.2. Validation interférométrique	155
5.6. Classement des hypothèses	157
5.6.1. Détermination de l’hypothèse référence.....	157
5.6.2. Analyse des différences de signatures.....	159
5.7. Résultats	159
5.8. Conclusion.....	167
6. EXTENSIONS.....	169
6.1. Introduction	169
6.2. Vers l’interférométrie polarimétrique	169
6.2.1. Notions de polarimétrie	170
6.2.2. L’interférométrie polarimétrique	172
6.2.3. Modèle de mélange de deux contributions ayant des mécanismes différents	173
6.2.4. Optimisation de cohérence	174
6.2.5. Estimation de la hauteur de bâtiments sur données interférométriques polarimétriques	174
6.2.6. Conclusion	177
6.3. Vers la très haute résolution	177
6.3.1. Influence de l’amélioration de la résolution.....	177
6.3.2. Présentation des données.....	178
6.3.3. Résultats de la méthode de reconstruction	179
6.3.4. Conclusion	184
6.4. Améliorations	184
6.4.1. Génération de contraintes sur les échos triples.....	184
6.4.2. Amélioration de la méthode de reconstruction.....	185
6.4.3. Conclusion.....	187
6.5. Conclusion.....	188
CONCLUSION GENERALE.....	189

PUBLICATIONS PERSONNELLES	193
BIBLIOGRAPHIE.....	195
ANNEXES	201
A. PRESENTATION DES DONNEES RAMSES	203
B. ETUDE STATISTIQUE DES IMAGES EN AMPLITUDE RAMSES SUR DUNKERQUE.....	207
C. GENERALITES SUR LES CHAMPS DE MARKOV.....	215
D. RESULTATS COMPLEMENTAIRES SUR LE MEAN SHIFT EN IMAGERIE RSO	217
E. REGULARISATION PAR PROGRAMMATION DYNAMIQUE.....	221

Acronymes

DLR	Deutsches Zentrum für Luft- und Raumfahrt
ENVISAT	Environmental Satellite
ESA	European Space Agency
FGAN	Forschungsinstitut für Angewandte Naturwissenschaften
FHR	Forschungsinstitut für Hochfrequenzphysik und Radartechnik
ONERA	Office National d'Etudes et de Recherches Aéronautiques
PAMIR	Phase Array Multifunctional Imaging Radar
RAMSES	Radar Aéroporté Multi-Spectral d'Etude des Signatures
RSO	Radar à Synthèse d'Ouverture
RADAR	Ration detection and ranging

Notations géométriques

B	base formée par les deux antennes RADAR
B_{ortho}	base orthogonale formée par les deux antennes RADAR
β	angle entre l'axe de la trajectoire et un bâtiment de la scène
D	distance capteur-cible en géométrie RADAR
d	distance entre cibles en géométrie RADAR
H	hauteur du capteur par rapport sol
h	hauteur d'un objet par rapport au sol
h_{amb}	hauteur d'ambiguïté
L, l	longueur et largeur de l'antenne physique
y	azimut en géométrie RADAR

Notations statistiques

a	amplitude de l'image RSO
h_A, h_C, h_S	paramètres de bande amplitude, cohérence, spatiale du filtre mean shift
i	intensité de l'image RSO
K	noyau du filtre mean shift
k	profil du noyau K
L	nombre de vues
M, μ	paramètres statistiques des lois de Rayleigh-Nakagami, Fisher, K
p_d, p_{fa}	probabilité de détection et de fausse alarme
R	réflectivité
σ_φ, σ_h	écart-type sur une mesure interférométrique et sur une mesure de hauteur

Notations RSO

α	ouverture angulaire de l'antenne RADAR
B_u	bande utile du chirp émis
c	vitesse de la lumière
ζ	angle de squint
F_{dop}	fréquence doppler
γ	cohérence empirique
λ	longueur d'onde

L_{synth}	longueur de l'antenne synthétique
pas_azi, pas_dist	pas d'échantillonnage selon l'axe azimut et l'axe distance
φ	phase de l'image RSO
r_{az}, r_{dist}	résolutions de l'image RADAR en azimut et en distance
res_{az}, res_{dist}	résolutions de l'image RSO en azimut et en distance
σ	surface équivalente RADAR
τ	durée du train d'onde émis par le capteur RADAR
θ	angle d'incidence
T_{int}	temps d'intégration
v	vitesse du porteur du capteur RADAR

Introduction générale

Les grosses agglomérations (villes, banlieues, zones industrielles) concentrent une grande partie de l'activité humaine. Pour observer et analyser cette activité et ses conséquences sur l'environnement par exemple, la télédétection, technique permettant de collecter des données sur un objet sans contact direct entre l'objet et le capteur, semble un moyen particulièrement adapté. De nombreux thèmes de recherche sont associés aux villes : la détection de changements (expansion des villes, destruction de bâtiments etc...), la surveillance du trafic, la mesure de pollution, la cartographie... Pour effectuer ces tâches, différents types de capteurs, entre autres imageurs, utilisant différentes parties du spectre des fréquences sont disponibles.

La technique d'imagerie la plus ancienne est l'imagerie optique utilisant le spectre du visible (longueur d'onde de 400nm à 800nm). Dans ce cas, le capteur est passif et capte les photons réfléchis par la scène. Ce type d'image est, la plupart du temps, facile à interpréter mais ne peut être obtenu que le jour et par beau temps.

L'imagerie RADAR à synthèse d'ouverture (RSO) propose une alternative intéressante à l'imagerie optique car elle utilise un capteur actif dans le spectre des ondes radio. Par conséquent, les images RSO peuvent être obtenues de jour comme de nuit, par temps nuageux ou pluvieux. Cette technique d'imagerie se révèle adaptée aux situations d'urgence (catastrophes naturelles, missions opérationnelles).

Les premiers capteurs RSO satellitaires mis en service furent SIR-A, SIR-B (années 80), ERS1 et RADARSAT (années 90). Ces derniers fournissaient des images basses résolutions (de l'ordre de la dizaine de mètres), adaptées à l'étude des reliefs. Avec l'avènement de capteurs haute résolution (de 20cm à quelques mètres) voire très haute résolution (inférieure à 20cm), les applications de reconnaissance de cibles ou bâtiments sont désormais possibles. Jusqu'à présent, les capteurs haute résolution étaient uniquement aéroportés et destinés à un usage militaire. Par conséquent, très peu d'informations sur ces systèmes d'acquisition sont disponibles. On peut néanmoins citer les capteurs RAMSES de l'ONERA ou PAMIR du FGAN-FHR qui permettent de former des images haute et très haute résolution. Avec le lancement de TerraSAR-X, capteur satellitaire du DLR, des images à 1m de résolution seront bientôt disponibles.

L'interprétation des images fournies par ces capteurs peut être effectuée manuellement au cas par cas, mais, devant l'augmentation croissante de la quantité de données à disposition, l'aide et/ou l'automatisation de l'interprétation est d'un intérêt majeur. Ce besoin est notamment très important en imagerie RSO où les images sont difficiles à interpréter.

Dans le cadre de cette thèse, nous nous intéressons au domaine de la reconstruction de bâtiments c'est-à-dire, à la représentation, sous forme d'objets géométriques en trois dimensions (3D) de bâtiments. Le thème de la reconstruction de bâtiments sur images RSO est relativement récent : les premiers travaux datent de la fin des années 90 [Kirscht 1998], [Gamba 1999,2000]. Les méthodes proposées récemment ([Soergel 2006], [Thiele 2006]) utilisent plusieurs acquisitions de la même scène à des résolutions comprises entre 20 et 50cm.

La détection de bâtiments sur images RSO est hors du cadre de cette thèse. Nous supposons que cette étape est déjà effectuée et la reconstruction est faite sur des imagerie comportant un bâtiment isolé. Le but de cette thèse est de proposer une méthode de

reconstruction de bâtiments à partir d'une unique acquisition interférométrique en bande X. Cette méthode est testée sur des images monopasses, à 50cm de résolution, issues du capteur RAMSES sur la zone industrielle du port de Dunkerque.

Cette thèse est divisée en deux parties : la première étudie la reconstruction de bâtiments sur images RSO. La seconde propose une méthode de reconstruction de bâtiments testée sur images RAMSES.

La première partie consiste tout d'abord en une présentation de l'imagerie RSO interférométrique dans le cadre d'un capteur aéroporté haute résolution (chapitre 1). Contrairement à d'autres types d'images, la formation des images RSO ne consiste pas simplement en l'acquisition d'un signal reçu. Des techniques de traitement du signal sont nécessaires pour obtenir des images de résolution suffisante pour l'application voulue. C'est pourquoi l'imagerie RSO, nécessitant d'importantes puissances de calculs, est assez récente. Le principe d'interférométrie RSO est également présenté. L'interférométrie RSO est une technique d'imagerie RADAR apparue au milieu des années 70 et consistant à acquérir deux images de la même scène avec des axes de visées très proches. Elle permet d'obtenir une information altimétrique dense sur la scène. Pour les applications nécessitant un aspect 3D, l'interférométrie semble être une technique intéressante.

Le deuxième chapitre de cette partie est dédié à l'analyse et l'interprétation des images RSO. Nous nous intéressons à l'origine des particularités des images RSO qui rendent l'interprétation difficile. Tout d'abord, elles peuvent être d'origine géométrique, provenant de l'acquisition en distance du capteur. Le repliement, qui apparaît lorsque deux zones d'altitudes différentes sont à la même distance du capteur, en est un exemple. Ensuite, la phénoménologie RADAR en milieu urbain est complexe car l'onde RADAR peut subir de multiples réflexions. Les phénomènes de réflexions simples, doubles et triples sont analysés sur données interférométriques. Enfin un modèle de mélange altimétrique de deux contributions dans les zones de repliement est proposé. Grâce à ce modèle, connaissant le mélange et l'une des deux contributions, nous montrons qu'il est possible d'obtenir la seconde. Ce modèle est ensuite validé sur l'estimation de la hauteur interférométrique de bâtiments.

Le troisième chapitre est dédié à l'étude bibliographique sur la représentation des bâtiments en 3D à partir d'images RSO. Les diverses méthodes de reconstruction sont regroupées selon les approches choisies, à savoir l'approche surfacique, n'effectuant aucune hypothèse sur la zone à reconstruire et aboutissant à un modèle numérique de surface (MNS), et l'approche objet : il s'agit de l'étape de reconstruction permettant d'obtenir un modèle numérique d'élévation (MNE). L'approche objet peut, elle-même, se décomposer selon une approche ascendante (on part de l'image pour obtenir un modèle), ou descendante (on part d'un modèle dont on détermine les paramètres). Cette étude bibliographique nous a permis de définir l'approche choisie.

La deuxième partie de cette thèse propose une chaîne de reconstruction suivant une approche objet ascendante, pour des modèles de bâtiments ayant une empreinte à angles droits et à toits plats ou plats par morceaux (en escalier) à partir d'une unique acquisition interférométrique de la scène. Cette méthode se décompose en deux étapes : l'extraction des signatures puis la reconstruction.

Le chapitre 4 est dédié à l'extraction des signatures de bâtiments sur images RAMSES. Pour cela, une étape de préfiltrage est utilisée pour réduire le chatoiement. Les filtres usuels en imagerie RSO sont paramétriques, c'est-à-dire qu'ils nécessitent un *a-priori* sur les statistiques au sein des images RSO. Nous proposons dans ce chapitre d'introduire le filtre

non-paramétrique mean shift ([Fukanaga 1975], [Commaniciu 2002]) en imagerie RSO. Ce filtre est tout d'abord comparé aux filtres RSO usuels sur images RSO synthétiques puis sur images RAMSES. Ensuite, le domaine de filtrage du mean shift est étendu à l'image de cohérence pour améliorer la détection des ombres. La fin de ce chapitre décrit les algorithmes de détection et d'extraction des signatures de bâtiments. Ils utilisent les hypothèses choisies sur les bâtiments à reconstruire et sont basés principalement sur la transformée de Hough.

Le chapitre 5 décrit l'algorithme de gestion d'hypothèses pour la reconstruction de bâtiments. A partir d'un ensemble incomplet de primitives, des hypothèses de bâtiments sont générées de façon à satisfaire les modèles de bâtiments choisis. Ensuite, les hypothèses obtenues sont validées ou rejetées grâce à une série de règles (géométriques...). Les hypothèses validées sont finalement classées pour aider l'utilisateur à effectuer la décision finale.

Enfin, le chapitre 6 décrit les extensions de ce travail de thèse montrant que les contributions proposées dans les chapitres précédents sont valides sur d'autres images. Tout d'abord le modèle de séparation de mélange de deux contributions, proposé au chapitre 2, est testé sur un jeu d'images interférométriques polarimétriques haute résolution. Il est également comparé à un modèle de mélange polarimétrique proposé dans [Colin 2005] sur l'estimation de la hauteur de bâtiments. Puis l'algorithme de reconstruction est testé sur des images de meilleure résolution que Dunkerque mais restant dans le domaine de la haute résolution. Enfin des perspectives pour améliorer la méthode de reconstruction sont dégagées.

En conclusion, sont rappelées les contributions et résultats issus de ce travail de thèse.

**Partie 1 : Reconstruction de bâtiment en
imagerie RADAR à synthèse d'ouverture
interférométrique haute résolution**

1. L'imagerie RADAR à synthèse d'ouverture

1.1. Introduction

Le principe de l'imagerie RADAR à synthèse d'ouverture (RSO) s'appuie sur des concepts issus de l'électromagnétisme et du traitement du signal. Le capteur RADAR est un capteur actif. Il permet de prendre des images de jour comme de nuit. L'utilisation de longueurs d'onde centimétriques le rend insensible aux conditions météorologiques (nuages). C'est pourquoi, dans un contexte d'urgence militaire, l'imagerie RSO permet d'obtenir des images rapidement.

Les capteurs aéroportés, tels que RAMSES (ONERA) ou PAMIR (FGAN-FHR), fournissent des images très haute résolution (jusqu'à 8cm) permettant d'envisager des problématiques de reconnaissance de cibles. Les applications de l'imagerie RSO ne sont pas uniquement militaires. Par exemple, ENVISAT lancé en 2002 par l'ESA (Agence Spatiale Européenne), est un satellite civil possédant un capteur RADAR basse résolution (25m) dédié à l'observation de la Terre, de l'atmosphère, des terres et de la glace. Les futurs satellites (TerraSAR-X, RADARSAT-2, Cosmo-SkyMed) permettront d'obtenir des images à des résolutions métriques.

Dans ce chapitre, les principes de l'imagerie RSO et de l'interférométrie RSO, technique permettant d'obtenir une carte altimétrique dense de la scène, sont présentés dans le cas d'un capteur aéroporté.

1.2. L'imagerie RSO

Dans un premier temps, le principe de l'imagerie RADAR est détaillé. L'imagerie RSO est ensuite présentée comme une extension de l'imagerie RADAR. Les résolutions des images RADAR sont améliorées par la compression en distance et la synthèse d'ouverture. Ce chapitre reprend les notions fondamentales en imagerie RSO et s'inspire des ouvrages et articles [Maître 2001], [Henry 2003] et [CCRS 2004].

1.2.1. Les ondes RADAR

Les ondes RADAR ont un spectre de fréquences variant entre 30MHz et 300GHz. Ce spectre est découpé en différentes bandes qui ne sont pas toutes utilisées en imagerie RSO. Par exemple, l'atmosphère absorbe fortement les ondes de la Bande V (entre 50 et 75 GHz). D'autres bandes sont réservées à un usage militaire. Le tableau ci-dessous présente les bandes de fréquences les plus utilisées en imagerie RSO.

Bande	Fréquence (GHz)	Longueur d'onde (cm)	Bande	Fréquence (GHz)	Longueur d'onde (cm)
P	0.22-0.39	76.9-133	X	5.75-10.90	2.7-5.2
L	0.39-1.55	19.3-76.9	Ku	10.90-22	1.36-2.7
S	1.55-4.20	7.1-19.3	Ka	22-36	0.83-1.36
C	4.20-5.75	5.2-7.1	W	75-110	0.27-0.40

1.2.2. RADAR et image complexe

Le terme RADAR signifie « radio detection and ranging ». Le système RADAR se compose d'une ou plusieurs antennes fonctionnant en émission et/ou réception. L'antenne du système RADAR émet un train d'ondes cohérentes électromagnétiques qui peut être réfléchi par des cibles à proximité. La mesure du temps entre l'émission (t_E) et la réception (t_R) du train d'ondes permet de déterminer la distance D entre la cible et le capteur par :

$$D = \frac{t_R - t_E}{2c}$$

avec c la vitesse de la lumière supposée constante.

Un système est dit « imageur » lorsque l'information donnée par le capteur présente une indication sur la localisation en deux dimensions de la cible. Dans l'exemple de la figure 1, il s'agit du couple (distance, azimuth) où l'azimut représente l'angle entre une référence et la cible, le capteur étant orientable autour d'un axe.

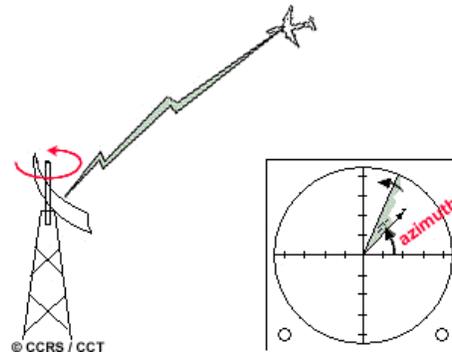


Figure 1 – système RADAR imageur au sol [CCRS 2004]

Dans le cas d'un capteur RADAR imageur satellitaire ou aéroporté, le principe est identique mais c'est le capteur qui se déplace. La trajectoire du porteur définit l'axe azimuth de l'image, noté y (cf. figure 2). La zone au sol, illuminée par le capteur, considérée selon cet axe est appelée trace. L'acquisition se fait généralement en visée latérale avec un angle d'émission θ . Cet axe de visée définit l'axe des distances noté d . La fauchée correspond à la zone illuminée entre la distance minimale (encore appelée distance proximale) et la distance maximale. Les deux axes azimuth / distance définissent les axes d'un repère cylindrique.

L'angle d'incidence est généralement non nul. En effet, lorsque l'angle d'incidence est nul, l'onde RADAR étant sphérique, deux zones du sol, situées à gauche et à droite du porteur, rétrodiffuseraient l'onde en même temps. Il serait alors impossible de faire la distinction entre ces ondes rétrodiffusées.

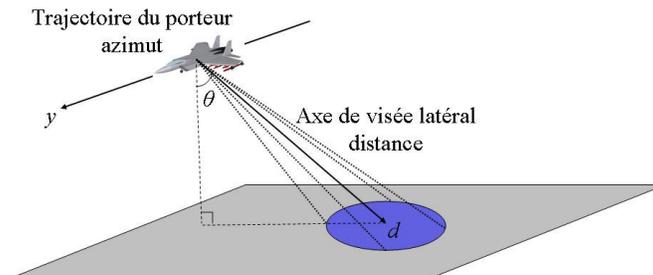


Figure 2 – système RADAR imageur aéroporté

L'utilisation d'une onde électromagnétique cohérente permet d'accéder à deux types d'informations : l'amplitude et la phase. Les images RSO se présentent donc sous la forme d'images complexes. La phase d'une onde en fonction de la distance D parcourue est :

$$\varphi = \frac{2\pi D}{\lambda} [2\pi] \quad (1)$$

avec λ longueur d'onde du signal RADAR.

Dès que l'onde RADAR a parcouru une distance supérieure à la longueur d'onde, l'information de distance ne peut plus être retrouvée à partir de la phase seule. De plus, à l'intérieur de la cellule de résolution, l'onde peut être réfléchiées par un ou plusieurs éléments placés aléatoirement. La phase des images RSO devient alors aléatoire. Néanmoins, pour certains types de cibles, des phénomènes de corrélation locale apparaissent parfois [Petit 2001].

1.2.3. Définition des résolutions en distance et en azimut

La résolution d'un système se définit par la distance minimale entre deux sources identiques placées à la distance D de l'observateur pour que ce dernier distingue les deux sources.

Considérons une antenne rectangulaire. En bande X, la longueur d'onde est de 3cm environ. La taille de l'antenne RADAR est de l'ordre de quelques longueurs d'onde. Soient L et l respectivement la longueur et la largeur de l'antenne. Sur la figure 3 est représenté le plan (P) orthogonal à l'axe des distances et situé à une distance D du capteur. Dans ce plan, le point Q a les coordonnées x et y selon les axes (Ox) et (Oy) .

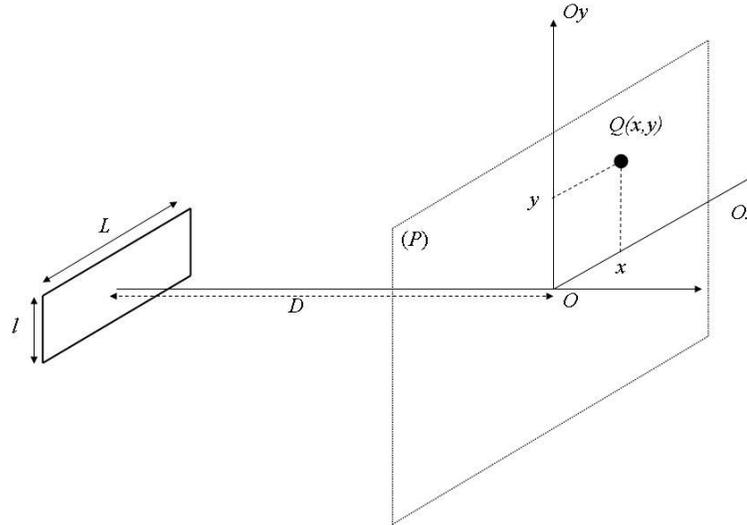


Figure 3 – Exemple d'antenne RADAR

Le RADAR imageur illumine des zones situées à plusieurs kilomètres. L'approximation de Fraunhofer est donc valide et le module du champ électromagnétique U en Q s'exprime sous la forme d'un produit de deux sinus cardinaux :

$$U(x, y) \propto \text{sinc}\left(\frac{\pi lx}{\lambda D}\right) \text{sinc}\left(\frac{\pi Ly}{\lambda D}\right) \quad (2)$$

avec : λ longueur d'onde du signal RADAR

L'ouverture angulaire de l'antenne à -3dB est notée α_{-3dB} et s'exprime par : $\sin(\alpha_{-3dB}) = 0.88\lambda/L$. Lorsque α_{-3dB} est petit : $\alpha_{-3dB} \approx 0.88\lambda/L$.

Sur la figure 4 sont représentées les énergies reçues par l'antenne rectangulaire sous forme de sinus cardinaux par deux sources distantes de d_0 , pour d_0 grand, d_0 à -3dB et d_0 faible. Lorsque d_0 diminue, la séparation des sources devient difficile voir impossible. La limite de bonne résolution étant subjective, dans la suite de ce document elle sera considérée comme la largeur du sinus cardinal à -3dB.

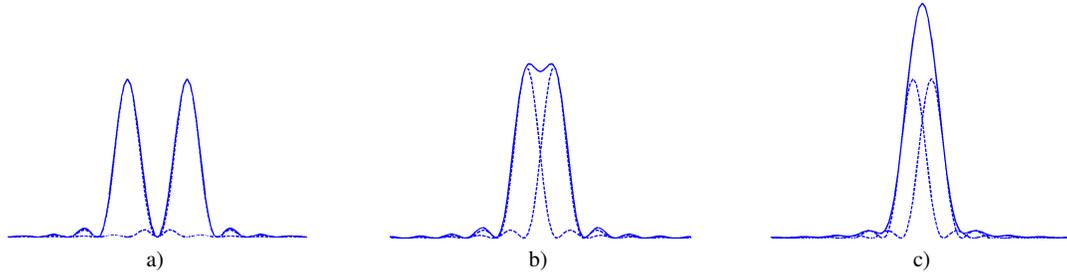


Figure 4 – Illustration de la résolution avec a) d_0 grand, b) $d_0 =$ largeur à -3dB et c) d_0 faible en pointillés sont représentées les énergies des sources prises séparément, en trait plein celle du mélange

En imagerie RADAR, deux résolutions sont définies : en distance et en azimut. La résolution en distance caractérise la capacité à séparer deux objets selon l'axe des distances. Elle dépend intrinsèquement de la durée τ du train d'ondes émis. Selon l'axe des distances, le train d'ondes occupe la distance $c\tau$ (cf. figure 5a). Les signaux réfléchis par des cibles sont séparables si la distance qui sépare les cibles est suffisamment grande pour que le capteur ait le temps de recevoir deux impulsions distinctes de durée τ . Soient deux cibles identiques 1 et 2, respectivement à la distance D_1 et D_2 du capteur (figure 5a).

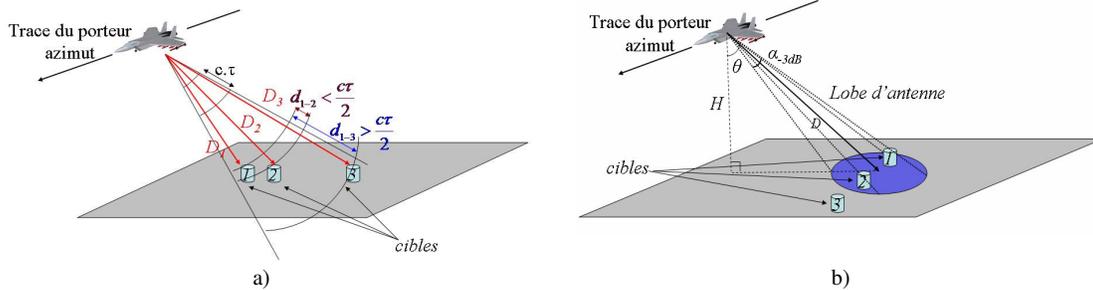


Figure 5 – Représentation de la résolution a) en distance et b) en azimut

Le capteur reçoit les trains d'ondes aux instants t_{R1} et t_{R2} . D'où :

$$t_{R2} - t_{R1} = \frac{2(D_2 - D_1)}{c} = \frac{2d_{1-2}}{c} < \tau$$

Les cibles 1 et 2 ne sont donc pas séparables. En revanche, les cibles 1 et 3, espacées de d_{1-3} , sont bien résolues. La résolution en distance est définie par :

$$r_{dist} = \frac{c\tau}{2} \quad (3)$$

Considérons le cas « classique » d'un capteur imageur aéroporté en bande X ($\lambda=3.17\text{cm}$) éclairant une zone située à 4km de distance à la vitesse de 80m/s. Pour obtenir une image à 50cm de résolution en distance, il faudrait pouvoir émettre un train d'onde de durée 3.3ns suffisamment énergétique pour parcourir le trajet aller-retour dans l'atmosphère. Techniquement, ce système n'est pas réalisable en imagerie RSO alors que c'est possible dans d'autres domaines comme l'imagerie ultrasonore.

Un ordre de grandeur de durée d'impulsion usuelle est de 20 μs . Ce qui donne une résolution en distance de 3km.

La résolution en azimut caractérise la capacité à séparer deux objets selon l'axe des azimuts. Selon cet axe, la résolution se formule à partir de sinus cardinaux (2) et en particulier par le terme $\text{sinc}\left(\frac{\pi Ly}{\lambda D}\right)$. En considérant la largeur du sinus cardinal à -3dB comme limite de séparation, la résolution à la distance D de l'antenne s'écrit par :

$$r_{az} = 0.88 \frac{\lambda}{L} D \quad (4)$$

Sur la figure 5b, deux cibles identiques à l'intérieur du cône d'angle solide α_{-3dB} (cibles 1 et 2) ne peuvent pas être résolues car les lobes principaux des sinus cardinaux sont trop proches. En revanche, les cibles 1 et 3 sont résolues

Dans le cas pratique étudié précédemment, pour obtenir une résolution de 50cm en azimut, il faudrait une antenne de 220m, beaucoup trop grande pour être montée sur un avion. Par exemple, le Transall, porteur du capteur RAMSES (cf. annexe A), mesure 32m de longueur environ. En supposant la taille de l'antenne égale à quatre fois la longueur d'onde, la résolution en azimut est de 800m.

1.2.4. Les méthodes d'amélioration de la résolution

Des contraintes de réalisations ne permettent pas d'obtenir des images RADAR avec des résolutions sub-métriques. Des techniques de traitement du signal sont alors utilisées pour améliorer ces résolutions : il s'agit de la compression d'impulsion pour la direction en distance et de la synthèse d'ouverture pour la direction azimutale.

1.2.4.1. Compression en distance

Pour améliorer la résolution en distance, il est possible d'effectuer une compression dès lors que l'onde électromagnétique émise est modulée par un « chirp » :

$$\begin{cases} A(t) = A_0 \exp\left(2j\pi\left[f_0 t + \frac{Kt^2}{2}\right]\right) & \text{si } |t| \leq \frac{\tau}{2} \\ A(t) = 0 & \text{sinon} \end{cases}$$

f_0 est la fréquence centrale de la porteuse. La fréquence instantanée f (dérivée de la phase) vue par la cible est alors :

$$f = f_0 + Kt$$

et correspond à une rampe dans le domaine fréquentiel. La largeur de bande du chirp, appelée bande utile, est notée B_u , et vaut $K\tau$.

La réponse fréquentielle de ce signal peut s'approcher par :

$$\begin{cases} \hat{A}(f) = \frac{1}{\sqrt{K}} \exp\left(j\frac{\pi}{4}\right) \exp\left(j\pi \frac{(f - f_0)^2}{K}\right) & \text{si } |f - f_0| \leq \frac{B_u}{2} \\ \hat{A}(f) = 0 & \text{sinon} \end{cases} \quad (5)$$

Après filtrage adapté du signal reçu par un filtre de réponse fréquentielle $\hat{A}^*(f)$ l'expression du signal détecté est la suivante :

$$g(t) = \tau \text{sinc}(B_u \cdot \pi \cdot t) \exp(-j2\pi f_0 t)$$

La résolution du système après compression est donnée par la largeur à -3dB du sinus cardinal détecté et ne dépend plus que de la largeur de bande :

$$res_{dist} = \frac{0.88.c}{2.B_u} \quad (6)$$

où res_{dist} est la résolution en distance après la compression. Le gain en résolution en distance est de $r_{dist} / res_{dist} = 0.88K\tau^2$.

Dans le cas étudié au §1.2.3, pour obtenir une résolution en distance de 50cm, il faut utiliser une bande utile de 260MHz. Le gain en résolution est de 6000 environ. La modulation en fréquence a permis d'utiliser un signal de durée initiale 20µs comme un signal de durée 3,3ns.

1.2.4.2. La synthèse d'ouverture

1.2.4.2.1. L'angle de squint

L'angle de squint est défini par l'angle entre l'axe de visée (axe des distances) du capteur et le vecteur orthogonal au vecteur vitesse, noté \vec{v} . Sur la figure 6, en rouge et en bleu sont représentés les axes de visées pour un angle de squint respectivement nul et non-nul.

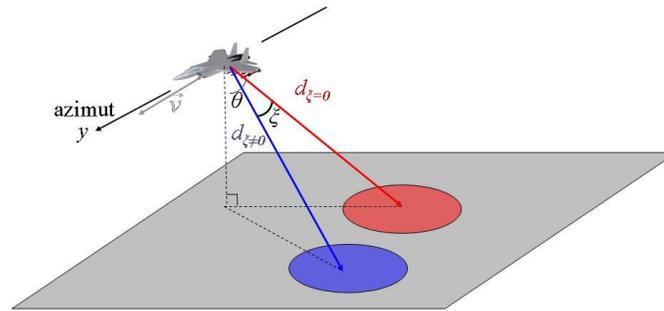


Figure 6 – angle de squint

1.2.4.2.2. La synthèse d'ouverture avec un angle de squint non-nul

Dans le cas général, l'angle de squint est non-nul. Pour améliorer la résolution en azimuth du système imageur, la nature cohérente de l'onde électromagnétique est utilisée pour synthétiser une antenne de dimension plus importante. Chaque zone est illuminée plusieurs fois par la même antenne à des azimuts différents (cf. figure 7). La longueur de l'antenne synthétique et le temps d'intégration sont notés L_{synth} et T_{int} . La trajectoire du porteur est supposée rectiligne uniforme à la vitesse v .

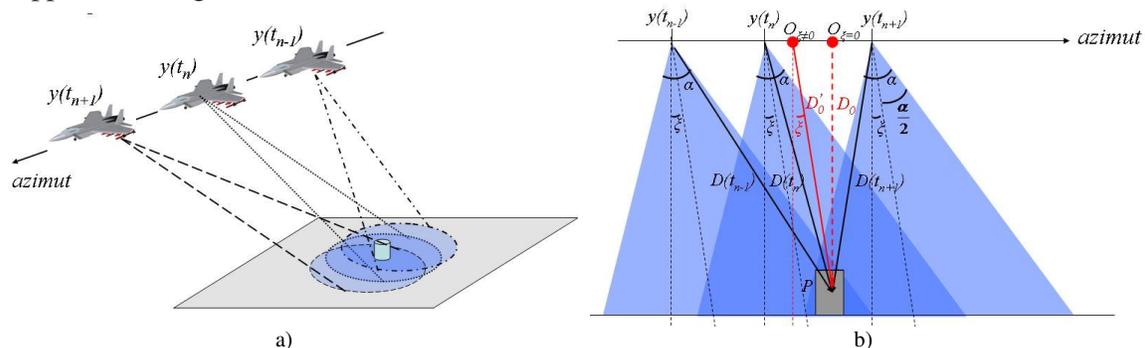


Figure 7 – Principe de la synthèse d'ouverture a) vue de l'avion, b) vue du sol aux instants t_{n-1} , t_n et t_{n+1}

Soient $O_{\xi \neq 0}$ et $O_{\xi=0}$ les points de la trajectoire situés à la distance minimum de la cible lorsque l'angle de squint est respectivement non-nul et égal à 0. Les distances $O_{\xi \neq 0}P$ et $O_{\xi=0}P$ sont notées D'_0 et D_0 .

En imagerie RSO aéroportée, la synthèse d'ouverture n'est pas effectuée sur toute l'ouverture angulaire de l'antenne, notée $\alpha_{.3dB}$, cf. (4), mais sur une partie seulement, notée α (avec $\alpha < \alpha_{.3dB}$), afin d'obtenir une résolution en azimut égale à celle en distance.

La longueur de l'antenne synthétique à l'issue du temps d'intégration T_{int} est :

$$L_{synth} = D'_0 \cos(\xi) \left(\tan\left(\frac{\alpha}{2} - \xi\right) + \tan\left(\frac{\alpha}{2} + \xi\right) \right) \quad (7)$$

La distance entre le porteur et la cible varie à chaque impulsion :

$$D(t) = D'_0 \sqrt{1 + \frac{y^2(t)}{D_0'^2} - \frac{2y(t)}{D'_0} \sin(\xi)} \quad (8)$$

En considérant le 0 de la trajectoire en $O_{\xi \neq 0}$, la distance minimale entre le capteur et la cible vaut D'_0 est obtenue lorsque :

$$y(t) = D'_0 \sin(\xi)$$

La relation (8) peut être approchée par :

$$D(t) \approx D'_0 \left(1 + \frac{y^2(t)}{2D_0'^2} \cos^2(\xi) - \frac{y(t)}{D'_0} \sin(\xi) \right) \quad (9)$$

Le terme en cosinus correspond à la courbure de l'hyperbole centrée en $O_{\xi \neq 0}$.

$\varphi(t)$, la phase instantanée de l'onde électromagnétique peut s'exprimer en fonction de la distance $D(t)$ parcourue par l'onde à l'instant t et de la longueur d'onde λ du signal RADAR par :

$$\varphi(t) = \frac{4\pi}{\lambda} D(t) \approx \frac{4\pi}{\lambda} D'_0 + \frac{2\pi \cos^2(\xi)}{\lambda D_0'} y^2(t) - \frac{4\pi \sin(\xi)}{\lambda} y(t)$$

L'expression de la fréquence Doppler est donnée par :

$$F_{dop}(t) = \frac{1}{2\pi} \frac{d\varphi(t)}{dt}$$

$$F_{dop}(t) = \frac{1}{2\pi} \frac{d\varphi(t)}{dy(t)} \frac{dy(t)}{dt}$$

La trajectoire du porteur est rectiligne uniforme donc l'azimut s'exprime par :

$$y(t) = -vt$$

D'où l'expression de la fréquence Doppler instantanée :

$$F_{dop}(t) = -\frac{2v \cos^2(\xi)}{\lambda D_0'} y(t) + \frac{2v \sin(\xi)}{\lambda}$$

$$F_{dop}(t) = \frac{2v^2 \cos^2(\xi)}{\lambda D_0'} t + \frac{2v \sin(\xi)}{\lambda} \quad (10)$$

Le premier terme de l'équation (10) varie linéairement en fonction de t et est à l'origine de la variation de la fréquence doppler. La fréquence doppler est positive lorsque le porteur se rapproche de $O_{\xi \neq 0}$ et négative lorsqu'il s'en éloigne. Le deuxième terme de

l'équation (10) correspond au doppler centroïde : le spectre de l'image selon l'axe doppler est décalé autour de cette valeur et n'est plus centré en 0.

La bande doppler totale utilisée lors de l'intégration s'exprime par :

$$B_{dop} = \frac{2v^2 \cos^2(\xi)}{\lambda D'_0} T_{int}$$

De plus :

$$T_{int} = \frac{L_{synth}}{v} \quad (11)$$

D'où l'expression de la bande doppler en fonction de l'ouverture angulaire :

$$B_{dop} = \frac{2v \cos^3(\xi)}{\lambda} \left(\tan\left(\frac{\alpha}{2} - \xi\right) + \tan\left(\frac{\alpha}{2} + \xi\right) \right) \quad (12)$$

La résolution à -3dB en azimuth s'exprime par :

$$res_{az} = 0.88 \frac{dy}{dF_{dop}} \delta F_{dop} = 0.88 \frac{dy}{dt} \frac{dt}{dF_{dop}} \delta F_{dop} \quad (13)$$

où δF_{dop} représente la résolution doppler. Elle est reliée au temps d'intégration par :

$$\delta F_{dop} = \frac{1}{T_{int}} \quad (14)$$

De plus, en dérivant (10) par rapport à t on obtient :

$$\begin{aligned} \frac{dF_{dop}}{dt} &= \frac{2v^2 \cos^2(\xi)}{\lambda D_0^2} \\ \frac{dF_{dop}}{dt} &= \frac{B_{dop}}{T_{int}} \end{aligned} \quad (15)$$

En combinant les relations (13), (14) et (15) l'expression de la résolution en azimuth est :

$$res_{az} = \frac{0.88v}{B_{dop}}$$

En remplaçant l'expression de B_{dop} par (12), on a :

$$res_{az} = \frac{0.88\lambda}{2 \cos^3(\xi) \left(\tan\left(\frac{\alpha}{2} - \xi\right) + \tan\left(\frac{\alpha}{2} + \xi\right) \right)} \quad (16)$$

1.2.4.2.3. La synthèse d'ouverture avec un angle de squint nul

Pour un angle de squint nul, les équations précédentes se simplifient. La longueur de l'antenne synthétique devient :

$$L_{synth} = 2D_0 \tan\left(\frac{\alpha}{2}\right) \quad (17)$$

La bande doppler et la résolution à -3dB s'écrivent alors :

$$B_{dop} = \frac{4v}{\lambda} \tan\left(\frac{\alpha}{2}\right) \quad (18)$$

$$res_{az} = \frac{0.88\lambda}{4 \tan\left(\frac{\alpha}{2}\right)} \quad (19)$$

1.2.4.2.4. La synthèse d'ouverture avec un angle de squint nul et une faible ouverture angulaire

Pour de faibles ouvertures angulaires, $\tan\left(\frac{\alpha}{2}\right) \approx \frac{\alpha}{2}$. D'où l'expression de la longueur de l'antenne synthétique :

$$L_{synth} = D_0\alpha \quad (20)$$

La bande et la résolution à -3dB en azimut s'écrivent :

$$B_{dop} = \frac{2v\alpha}{\lambda} \quad (21)$$

$$res_{az} = \frac{0.88\lambda}{2\alpha} \quad (22)$$

La résolution en azimut ne dépend plus de la distance entre le capteur et la cible mais de l'ouverture angulaire considérée. En effet, les relations (11) et (20) montrent qu'à ouverture angulaire et vitesse constante, les zones éloignées sont intégrées plus longtemps que les zones proches.

Lorsque le signal est intégré sur l'ouverture complète de l'antenne (supposée faible), la résolution s'améliore et devient :

$$res_{az} = \frac{0.88\lambda}{2\alpha_{-3dB}} = \frac{L}{2} \quad (23)$$

Pour avoir une résolution en azimut de 50cm, cela revient à considérer une ouverture angulaire pour l'antenne synthétique de 1.59° et une antenne physique plus petite que 1m. L'hypothèse de faible ouverture est donc valide. Pour une vitesse de 80m/s et une distance de 4000m, le temps d'intégration est de 1.38s. Le gain en résolution est de 1800 environ.

1.2.4.2.5. La compensation de mouvement

Pratiquement, dans le cas aéroporté, l'hypothèse d'un mouvement rectiligne et uniforme n'est jamais vérifiée. Un décalage existe entre la trajectoire rectiligne uniforme, encore appelée trajectoire nominale, et la trajectoire réelle. L'étape de compensation de mouvement consiste à transposer la synthèse d'ouverture de la trajectoire nominale à la trajectoire réelle.

Sur la figure 8, les trajectoires nominales et réelles du porteur sont représentées respectivement en noir et en rouge. Dans le cas de la trajectoire nominale, les positions du capteur en azimut utilisées pour la synthèse d'ouverture sont numérotées de 0 à 7 en noir. La trajectoire étant rectiligne uniforme, la différence d'azimut entre chaque position successive est constante.

Pour compenser la trajectoire réelle (numérotée de 0 à 7 en rouge), il faut déterminer les écarts entre les deux trajectoires, pour chaque position i du capteur. Cette différence de distance introduit un déphasage de l'onde électromagnétique. La synthèse est alors effectuée à partir de la somme cohérente des ondes électromagnétiques compensées.

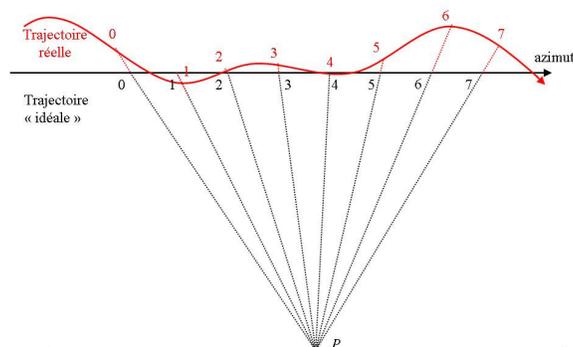


Figure 8 – Principe de la compensation de mouvement : les numéros indiquent les positions du porteur sélectionnées lors de la synthèse d'ouverture

1.2.5. Résolution et pas d'échantillonnage

La notion de résolution est une notion continue. En revanche, l'acquisition du signal RADAR est effectuée de façon discrète. Soient $F_{e,dist}$ et $F_{e,az}$ les fréquences d'échantillonnage minimales pour que le théorème de Shanon soit respecté.

En distance, la résolution est liée directement à la bande utile B_u (5). Cette plage de fréquence utilisée pour la compression en distance est centrée autour de la fréquence de la porteuse. La fréquence d'échantillonnage minimum est donc :

$$F_{e,dist} = 2B_u$$

RAMSES possède également un autre mode d'échantillonnage présenté en annexe A.

En azimut, la résolution dépend de l'ouverture angulaire (22). Dans le cas d'une acquisition à angle de squint nul et à faible ouverture angulaire, la fréquence maximale est de $\frac{2\alpha v}{0.88\lambda}$. La distance entre deux acquisitions successives définit la fréquence d'échantillonnage en azimut. La fréquence de répétition des impulsions (FRI) doit être supérieure à cette fréquence d'échantillonnage. Pour ne pas perdre d'information il faut :

$$FRI \geq F_{e,az} = \frac{4\alpha v}{0.88\lambda}$$

1.2.6. Conclusion

Les principes de l'imagerie RSO ont été présentés. La compression d'impulsion et la synthèse d'ouverture améliorent respectivement la résolution en distance et celle en azimut. Il est important de remarquer que ces résolutions ne dépendent plus de la distance entre le porteur et la cible mais simplement de la bande utilisée et de la taille de l'antenne synthétique. Les images RSO sont des images complexes. Les images de phases de deux images RSO, prises avec des angles d'incidence très proches, sont utilisées en interférométrie pour obtenir une carte dense d'altitude de la scène.

1.3. L'interférométrie RSO

1.3.1. Principe

Le principe de l'interférométrie n'est pas nouveau et n'est nullement réservé à l'imagerie RADAR. En effet, le premier interféromètre est apparu en 1881, inventé par l'américain Michelson. Le principe est applicable à tout type d'éclairement cohérent. En astronomie par exemple, les ondes radios sont utilisées pour améliorer la résolution des télescopes. Le radiohéliographe de Nançay utilise cette technique pour prendre des images du soleil [Radio 06].

En imagerie RSO, l'interférogramme est obtenu grâce à un couple d'images prises à des incidences très proches se mesurant en fraction de radians (cf. figure 9). L'acquisition des deux images peut se faire simultanément : elle est dite monopasse. Ce type d'acquisition est souvent utilisé avec les capteurs aéroportés. Le capteur émet sur une antenne et reçoit sur deux antennes. Pour les capteurs satellitaires, il s'agit en général du même capteur effectuant des mesures lors de passages sur le même site et sur la même orbite mais à des positions légèrement différentes. L'acquisition est multipasse et les prises de vue sont décalées dans le temps. Dans ce paragraphe, le cas d'une acquisition monopasse est considéré.

L'interférogramme monovue est obtenu par la différence des images de phase du couple interférométrique :

$$\varphi_0 = \arg(z_1 z_2^*) \quad (24)$$

où φ_0 interférogramme
 \arg argument du nombre complexe
 z_1 et z_2 valeurs complexes du couple interférométrique

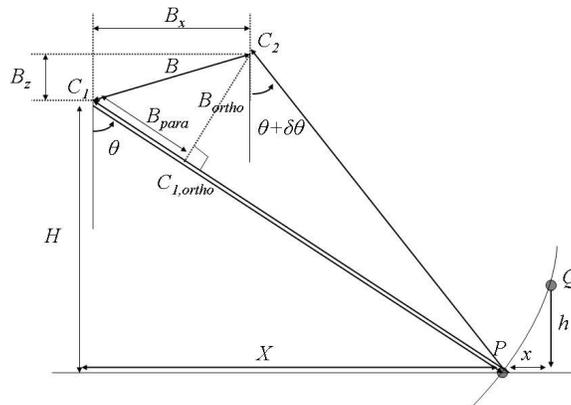


Figure 9 – Schéma de prise de vue d'un interférogramme

C_1 et C_2 sont les antennes utilisées pour prendre les deux images du couple interférométrique et sont appelées respectivement antenne maîtresse et antenne esclave. L'antenne C_1 est à une hauteur H du sol. La base interférométrique, notée B , est définie par les coordonnées de l'antenne esclave dans le référentiel lié à C_1 . La base interférométrique s'exprime par :

$$B \begin{pmatrix} B_x \\ B_z \end{pmatrix}$$

En géométrie terrain, dans le repère lié au sol, les coordonnées des antennes sont :

$$C_1 \begin{pmatrix} 0 \\ H \end{pmatrix} \quad C_2 \begin{pmatrix} B_x \\ H + B_z \end{pmatrix}$$

B_{ortho} est la base orthogonale, c'est à dire la composante orthogonale de la base sur la visée effectuée sous un angle d'incidence θ .

$$B_{ortho} = B_z \cos(\theta) + B_x \sin(\theta) \quad (25)$$

Le projeté de C_2 sur l'axe de visée est noté $C_{1,ortho}$. La distance $C_1 C_{1,ortho}$ définit la composante de la base parallèle à l'axe de visée. Cette composante est notée B_{para} .

Le point P est supposé d'altitude connue (nulle). Il sert de référence pour la mesure interférométrique. Q est un point dont l'altitude est inconnue appartenant au même front d'onde que P ($C_1 P = C_1 Q$). Dans le référentiel lié au sol, les coordonnées de ces deux points s'expriment par :

$$P \quad \begin{pmatrix} X \\ 0 \end{pmatrix} \quad Q \quad \begin{pmatrix} X + x \\ h \end{pmatrix}$$

En supposant le front d'onde plan (*hypothèse vérifiée au chapitre 2*) x et h vérifient :

$$h \cos(\theta) - x \sin(\theta) = 0 \quad (26)$$

1.3.2. Phase interférométrique et hauteur

$(C_1 P - C_2 P)$ représente la différence de trajet de l'onde RADAR entre l'antenne maître et le point de référence P et l'antenne esclave et ce même point. Par la relation (1), cette différence de parcours peut être reliée à une différence de phase appelée phase orbitale car elle est liée au trajet de l'onde dans l'atmosphère et non à la topographie de la zone imagée. Elle peut être considérée comme un offset dépendant de X .

De la même façon, $(C_1 Q - C_2 Q)$ peut être relié à une différence de phase appelée phase interférométrique qui contient le même offset de phase orbitale car P et Q sont à la même distance de C_1 . Après avoir retiré cet offset, il reste une phase appelée phase topographique, notée φ , liée à la position et à l'altitude de Q .

La différence de marche reliée à la phase topographique, notée δR , correspond donc à la différence entre les parcours $(C_1 Q - C_2 Q)$ et $(C_1 P - C_2 P)$, avec :

$$C_1 P - C_2 P = \sqrt{X^2 + H^2} - \sqrt{(X - B_x)^2 + (H + B_z)^2}$$

$$C_1 Q - C_2 Q = \sqrt{(X + x)^2 + (H - h)^2} - \sqrt{(X + x - B_x)^2 + (H - h + B_z)^2}$$

D'où :

$$\delta R = \sqrt{X^2 + H^2} \left(\sqrt{1 + \frac{x^2 + h^2 + 2Xx - 2Hh}{X^2 + H^2}} - \sqrt{1 + \frac{(x - B_x)^2 + (h - B_z)^2 + 2X(x - B_x) - 2H(h - B_z)}{X^2 + H^2}} \right) - 1 + \sqrt{1 + \frac{B_x^2 + B_z^2 - 2XB_x + 2HB_z}{X^2 + H^2}} \quad (27)$$

H^2 et X^2 sont très grand devant H , X , h , x , B_x et B_z . Par un développement limité à l'ordre 2 de l'expression (27) [Abdelfattah 98] et en ne gardant que les termes en X^2 , H^2 et $X.H$, la différence de marche s'exprime par :

$$\delta R = \frac{1}{(\sqrt{X^2 + H^2})^3} (XH (B_z x + B_x h) + H^2 B_x x + X^2 B_z h)$$

D'où :

$$\delta R = \frac{X^2 + H^2}{\left(\sqrt{X^2 + H^2}\right)^3} \left((B_z x + B_x h) \cos(\theta) \sin(\theta) + B_x x \cos^2(\theta) + B_z h \sin^2(\theta) \right)$$

$$\delta R = \frac{B_{ortho}}{\sqrt{X^2 + H^2}} (x \cos(\theta) + h \sin(\theta))$$

En utilisant (26), δR s'écrit :

$$\delta R = \frac{B_{ortho}}{\sqrt{X^2 + H^2}} \frac{h}{\sin(\theta)} \quad (28)$$

La relation entre la phase topographique et la différence de marche est :

$$\varphi = \frac{2\pi\delta R}{\lambda} \quad (29)$$

D'où :

$$\varphi = \frac{2\pi}{\lambda} \frac{B_{ortho}}{\sqrt{X^2 + H^2}} \frac{h}{\sin(\theta)} \quad (30)$$

Seule la composante orthogonale de la base intervient dans l'expression (30).

La phase est mesurée modulo 2π . La hauteur d'ambiguïté, notée h_{amb} , est introduite comme la hauteur pour laquelle la rotation de la phase des franges topographiques est égale à 2π :

$$h_{amb} = \frac{\lambda H \tan \theta}{B_{ortho}} \quad (31)$$

Une mesure de la hauteur h en un point de l'interférogramme s'exprime par :

$$h = \varphi \frac{h_{amb}}{2\pi} \quad [h_{amb}] \quad (32)$$

Ainsi, il n'est pas possible de faire la différence entre deux objets de hauteur h et $h+h_{amb}$.

L'écart-type altimétrique à partir de l'écart-type sur la phase est alors obtenu par :

$$\sigma_h = \frac{h_{amb} \cdot \sigma_\varphi}{2\pi} \quad (33)$$

Pour éluder le problème de l'ambiguïté de hauteur, le choix d'une faible base orthogonale semble naturel. Néanmoins, la précision sur la mesure interférométrique, qui est une précision sur la phase, limite ce choix. A bruit de phase constant, plus la base est faible et la hauteur d'ambiguïté est élevée, plus le bruit sur la hauteur interférométrique est important. Il faut donc faire un compromis sur B_{ortho} , dépendant de la tâche à effectuer. Pour de forts reliefs il faut une faible base orthogonale pour pouvoir dérouler les franges plus facilement. Le bruit sur la hauteur interférométrique est alors important. Pour de faibles reliefs, une grande base orthogonale est souhaitable pour avoir une meilleure précision sur la hauteur interférométrique.

1.3.3. Cohérence théorique

La matrice de cohérence ou covariance entre les deux images complexes du couple interférométrique s'écrit :

$$C_Z = \begin{bmatrix} E(|z_1|^2) & E(z_1 z_2^*) \\ E(z_1^* z_2) & E(|z_2|^2) \end{bmatrix}$$

Le coefficient de corrélation complexe ou degré de cohérence est défini par :

$$\rho_{12} = \frac{E(z_1 z_2^*)}{\sqrt{E(|z_1|^2)E(|z_2|^2)}} = \Gamma \exp(j\bar{\varphi})$$

où Γ représente la cohérence théorique et $\bar{\varphi}$ le déphasage effectif entre z_1 et z_2 . Par définition, Γ varie entre 0 et 1.

1.3.4. Interférogramme multivue et cohérence empirique

Pratiquement, la cohérence est calculée sur un nombre limité de cellules de résolution. Ce nombre est appelé nombre de vues ou « look » en anglais et est noté L . L'interférogramme multivue se calcule par :

$$\gamma \exp(j\varphi) = \frac{\sum_{k=1}^L z_{1,k} z_{2,k}^*}{\sqrt{\left(\sum_{k=1}^L z_{1,k} z_{1,k}^* \right) \left(\sum_{k=1}^L z_{2,k} z_{2,k}^* \right)}} \quad (34)$$

L'interférogramme multivue est la phase de ce produit normalisé. L'amplitude est appelée cohérence empirique. Pour $L=1$, l'expression de l'interférogramme complexe monovue est retrouvée et la cohérence empirique vaut 1.

La distribution de la cohérence empirique et de la phase en multivue s'écrivent [Maître 2001]:

$$p(\gamma|\Gamma, L) = 2(L-1)(1-\Gamma^2)^L \gamma(1-\gamma^2)^{L-2} {}_2F_1(L, L; 1; \gamma^2 \Gamma^2) \quad (35)$$

et

$$p(\varphi|\bar{\gamma}, L, \bar{\varphi}) = \frac{(1-\bar{\gamma}^2)^L}{2\pi} \frac{1}{2L+1} {}_2F_1\left(2, 2L; L + \frac{3}{2}; \frac{1+\bar{\gamma} \cos(\varphi - \bar{\varphi})}{2}\right) \quad (36)$$

avec $\bar{\gamma}$ et $\bar{\varphi}$ cohérence et phase moyenne de la zone

${}_2F_1$ la fonction hypergéométrique

Intuitivement, la cohérence empirique permet de savoir si les L cellules de résolution de la première image du couple interférométrique ont des « similitudes » de phase avec les mêmes L cellules de résolution de la seconde image. L'équation (34) permet de montrer que si les phases des deux images sont proches, la cohérence sera élevée. Lorsque ces phases sont différentes, la cohérence sera faible. La cohérence devient alors un indicateur de la qualité du signal interférométrique.

1.3.5. Bruits sur la mesure interférométrique

Les pertes de cohérence dues aux bruits thermiques et surfaciques sont notées $\gamma_{thermique}$ et $\gamma_{surfacique}$. Dans le cas d'une acquisition multipasse une perte de cohérence, notée $\gamma_{rotation}$, qui

dépend de l'angle formé par les deux trajectoires du porteur est introduite. En monopasse, ce terme n'intervient pas car les deux antennes du capteur sont solidaires et ont des trajectoires parallèles.

Le bruit thermique est un bruit additif inhérent au capteur. A chaque mesure effectuée, il modifie la phase du signal rétrodiffusé par la scène. Les deux images de phase du couple interférométrique sont bruitées indépendamment l'une de l'autre, d'où une perte de cohérence. La perte de cohérence due au bruit thermique s'exprime par [Bamler 1993] :

$$\gamma_{thermique} = \frac{1}{1 + RSB^{-1}}$$

Le bruit de phase surfacique provient de la différence d'angle d'incidence entre les deux images. En effet, localement la projection au sol de la fréquence de la porteuse du signal RADAR n'est pas la même. Le phénomène de 'wveshift' apparaît, se traduisant par un décalage fréquentiel en distance des spectres complexes des images. Les parties non communes du spectre n'apportent que du bruit. Ce décalage fréquentiel induit alors une perte de cohérence qui s'écrit, pour une fenêtre rectangulaire et un sol de pente ε [Gatelli 1994] :

$$\gamma_{surfacique} = \left(1 - \frac{\Delta f}{B_u}\right) \text{ où } \Delta f = \frac{c \cdot B_{ortho} \cdot \cos(\theta - \varepsilon)}{2 \cdot H \cdot \lambda \cdot \tan(\theta - \varepsilon)} \quad (37)$$

La notion de base orthogonale critique, notée $B_{ortho,c}$, est alors introduite et correspond à la valeur minimale de la base orthogonale pour laquelle la cohérence surfacique est nulle :

$$B_{ortho,c} = \frac{\lambda H \tan(\theta - \varepsilon) 2B_u}{\cos(\theta - \varepsilon) c}$$

La cohérence empirique mesurée est obtenue par :

$$\gamma = \gamma_{thermique} \gamma_{surfacique}$$

La précision altimétrique de la mesure interférométrique dépend du type d'interaction entre l'onde RADAR et la zone illuminée. Ces interactions peuvent être une rétrodiffusion surfacique de Rayleigh ou une réflexion spéculaire [Bamler 1993]. L'origine des différents comportements de l'onde RADAR en fonction des caractéristiques de l'onde incidente et de la zone éclairée est présentée en détail lors de l'analyse phénoménologique des images RSO aux §2.4.3 et §2.4.4.

Pour une rétrodiffusion surfacique de Rayleigh (cf. annexe B.2.1) sur une zone homogène, l'écart-type altimétrique est donné par :

$$\sigma_\varphi^2 = \int_{-\pi}^{\pi} \varphi^2 p(\varphi | \bar{\gamma}, L, \bar{\varphi}) \quad (38)$$

avec p la distribution de la phase donnée par l'équation (36).

Lorsqu'il y a un réflecteur dominant dans la cellule de résolution, la réflexion est spéculaire. γ est proche de 1 et l'écart-type interférométrique s'exprime par [Just 1994] :

$$\sigma_\varphi^2 = \frac{1 - |\gamma|^2}{2|\gamma|^2} \quad (39)$$

1.3.6. Conclusion

L'interférométrie RSO fournit deux types d'information. La première est l'interférogramme qui peut se voir comme une carte dense d'élévation. La seconde est la cohérence qui donne une information sur la qualité locale des données interférométriques en représentant les similarités des deux images de phase du couple interférométrique.

1.4. Conclusion

Les principes fondamentaux de l'imagerie RSO ont été présentés dans le cas de l'imagerie aéroportée. Les résolutions des images RSO ne dépendent plus de la distance entre le capteur et la zone illuminée mais des caractéristiques technologiques du capteur (Bande utile et longueur de l'antenne synthétique). De plus le concept d'interférométrie permet d'obtenir une information de phase dense qui peut être reliée à l'altitude de la scène ce qui est un atout majeur pour l'extraction de hauteur des éléments de la scène.

2. Etude et interprétation des images RSO haute résolution en milieu urbain

2.1. Introduction

Les images RSO sont difficiles à interpréter. En effet, le mode d'acquisition en visée latérale déforme géométriquement la scène. De plus l'utilisation d'une onde électromagnétique cohérente confère un aspect visuel bruité aux images RSO. Le milieu urbain complique également l'interprétation des images car les scènes urbaines présentent de nombreuses ruptures engendrant de multiples rebonds de l'onde électromagnétique.

2.2. Géométrie en imagerie RSO

2.2.1. Géométrie RADAR et géométrie terrain

Les images RSO sont échantillonnées selon l'axe des distances (cf. §1.2.5). Les images obtenues sont dites en géométrie RADAR. Il est possible d'effectuer une projection au sol des images en connaissant, pour chaque pixel, l'angle d'incidence local de l'onde électromagnétique. La taille d'un pixel projeté au sol n'est donc pas constante du début à la fin de la fauchée et dépend de sa position en distance. L'image ré-échantillonnée selon la longitude et la latitude est dite en géométrie terrain.

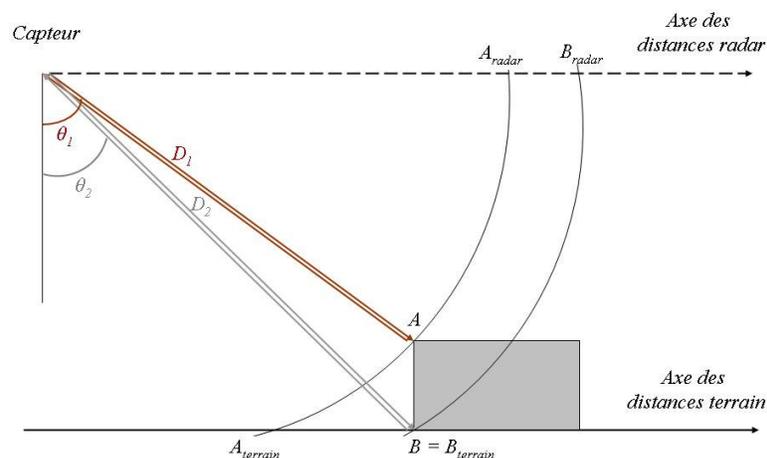


Figure 10 – Façade d'un bâtiment projeté en géométrie radar et en géométrie terrain

2.2.2. Hypothèse d'un front d'onde plan // front d'onde sphérique

L'onde émise par le capteur RADAR est sphérique. Pour une plus grande simplicité dans les calculs, le front d'onde est supposé plan. C'est-à-dire que l'incidence de l'onde électromagnétique peut être considérée comme constante localement. Cela revient à

confondre les angles θ_1 (angle d'incidence en haut de la façade) et θ_2 (angle d'incidence en bas de la façade) de la figure 10.

Si cette hypothèse est communément acceptée en imagerie RSO satellitaire, dans le cas aéroporté, elle doit être validée. Pour cela l'erreur entre les projections de la façade d'un bâtiment en géométrie RADAR avec et sans hypothèse d'onde plane, représentée sur la figure 11a, est calculée.

Soient :

h la hauteur du bâtiment

H la hauteur du capteur

X la position du bâtiment sur le sol en géométrie terrain

$\Delta R^{sphérique}$ la hauteur du bâtiment en géométrie RADAR sans hypothèse d'onde plane

ΔR^{plane} la hauteur du bâtiment en géométrie RADAR avec hypothèse d'onde plane

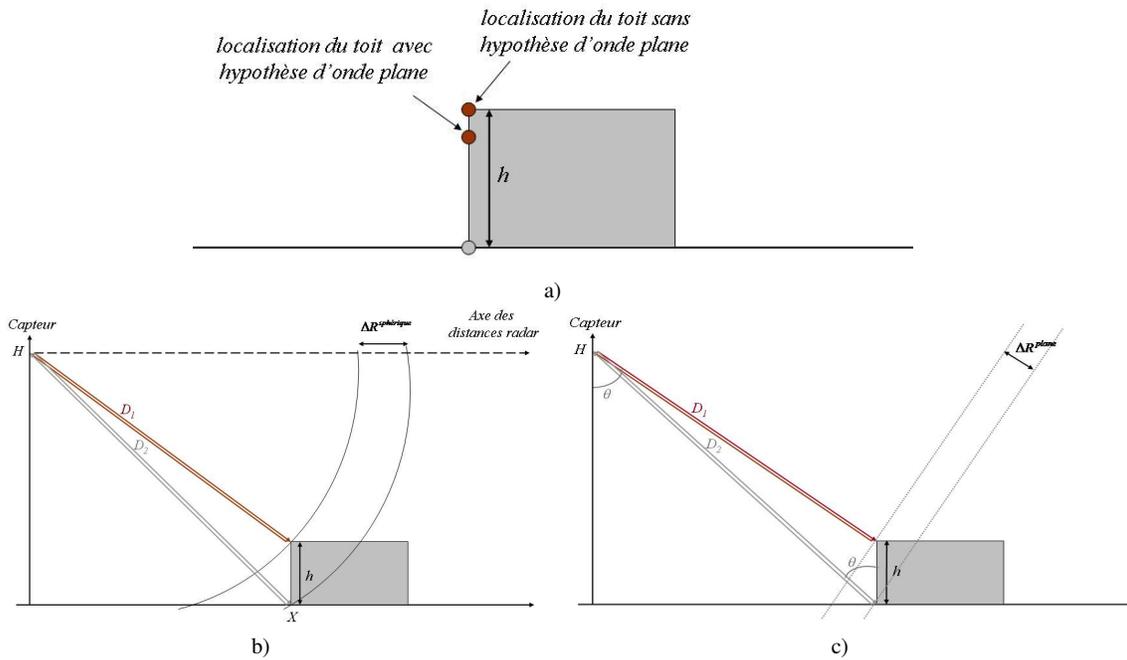


Figure 11 – a) erreur de localisation du toit avec l'hypothèse d'onde plane
Façade d'un bâtiment projetée en géométrie radar avec une onde b) sphérique et c) plane

Dans le cas d'une onde sphérique (schéma de la figure 11b), on a :

$$D_1^2 = (H - h)^2 + X^2$$

$$D_2^2 = H^2 + X^2$$

D'où

$$\begin{aligned} \Delta R^{sphérique} &= D_2 - D_1 = \sqrt{X^2 + H^2} - \sqrt{X^2 + (H - h)^2} \\ &= \sqrt{X^2 + H^2} - \sqrt{X^2 + H^2} \sqrt{\left(1 + \frac{h^2 - 2hH}{X^2 + H^2}\right)} \end{aligned}$$

L'ordre de grandeur de la hauteur du capteur et des distances sont, en Bande X aéroporté, respectivement de 3000m et de 4000m. La hauteur maximale des bâtiments

reconstruits est de l'ordre de 40-50m donc on peut effectuer un développement limité à l'ordre 2 de l'expression précédente :

$$\begin{aligned}\Delta R^{sphérique} &\approx \sqrt{X^2 + H^2} - \sqrt{X^2 + H^2} \left(1 + \frac{1}{2} \frac{h^2 - 2hH}{X^2 + H^2} - \frac{1}{8} \left(\frac{h^2 - 2hH}{X^2 + H^2} \right)^2 \right) \\ &\approx -\frac{1}{2} \frac{h^2 - 2hH}{\sqrt{X^2 + H^2}} + \frac{1}{8} \frac{(h^2 - 2hH)^2}{(X^2 + H^2)^{3/2}} \\ &\approx -\frac{1}{2} \frac{h^2 - 2hH}{D_2} + \frac{1}{8} \frac{(h^2 - 2hH)^2}{D_2^3}\end{aligned}$$

En négligeant les termes du troisième et quatrième ordre en h la hauteur du bâtiment s'écrit :

$$\Delta R^{sphérique} \approx \frac{hH}{D_2} + \frac{h^2(H^2 - D_2^2)}{2D_2^3} \quad (40)$$

Dans le cas d'une onde plane (schéma de la figure 11c) la hauteur du bâtiment en géométrie RADAR s'exprime par :

$$\Delta R^{plane} = h \cos(\theta) = \frac{hH}{D_2} \quad (41)$$

En utilisant les formules (40) et (41), l'erreur Err associée à l'hypothèse d'onde plane s'exprime par:

$$Err = \Delta R^{sphérique} - \Delta R^{plane} = \frac{1}{2} \frac{h^2}{D_2} \left(\frac{H^2}{D_2^2} - 1 \right)$$

De plus :

$$\frac{H^2}{D_2^2} - 1 = \frac{H^2 - D_2^2}{D_2^2} = \sin^2(\theta_2)$$

D'où

$$Err = \frac{h^2 \sin^2(\theta_2)}{2D_2} \quad (42)$$

La hauteur du capteur et la distance capteur-bâtiment sont respectivement de 3000m et 4000m. L'erreur Err est représentée pour un bâtiment de 40m de hauteur le long de l'axe des distances (les échelles sont en mètre) :

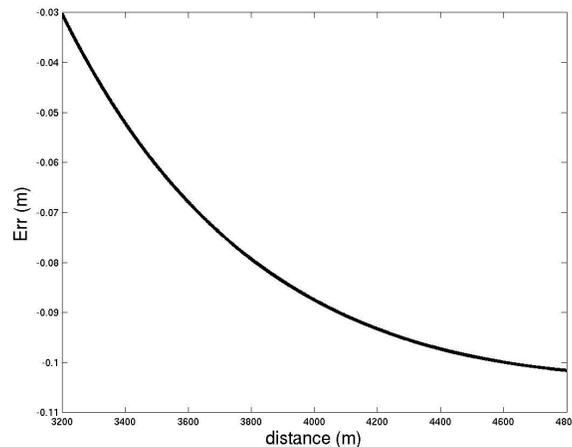


Figure 12 – Erreur sur le calcul de la hauteur d'un bâtiment en géométrie radar due à l'hypothèse d'onde plane

L'erreur sur la hauteur calculée à cause de l'hypothèse d'onde plane varie entre 3cm et 10cm. La résolution des images RAMSES en bande X sur le site de Dunkerque est de 50cm environ (cf. annexe A). L'erreur commise dans l'évaluation de la hauteur est donc inférieure à la résolution du capteur. L'hypothèse d'onde plane est donc valide. Il faut noter que pour une résolution de 10cm, cette approximation est à la limite de validité.

Nous nous intéressons maintenant à la parallaxe. Considérons le toit plat d'un bâtiment représenté en géométrie terrain sur la figure 13a. Les segments AB et CD sont parallèles en géométrie terrain. Dans quelle mesure le reste-t-il en géométrie RADAR ?

Soient D_x , Y_x et θ_x respectivement la distance entre capteur et le point x , l'azimut du point x et l'angle d'incidence de l'onde RADAR en x avec $x = \{A,B,C,D\}$. La distance D_A est choisie comme distance de référence. La hauteur du capteur par rapport au toit du bâtiment est notée H . La longueur, largeur et l'orientation du bâtiment sont notées respectivement L , l et β .

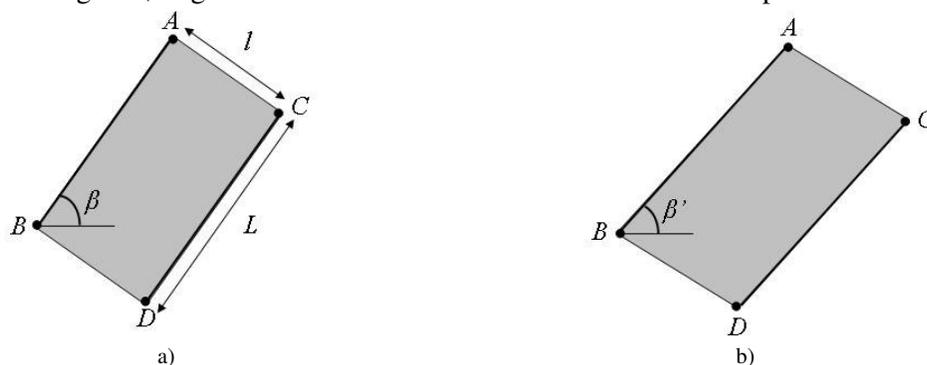


Figure 13 – représentation de deux segments parallèles en géométrie a) terrain et b) RADAR avec hypothèse d'onde plane

Les pentes $p_{AB}^{terrain}$ et $p_{CD}^{terrain}$ des segments AB et CD en géométrie terrain s'expriment en fonction de l'orientation β du bâtiment.

$$p_{AB}^{terrain} = p_{CD}^{terrain} = -\tan(\beta)$$

En géométrie RADAR, les pas d'échantillonnages selon la distance et l'azimut ne sont pas toujours identiques. Par conséquent, lorsque l'onde est considérée comme plane, les pentes des segments AB et CD, notées

$$p_{AB}^{RADAR} = p_{CD}^{RADAR} = -\frac{pas_azi}{pas_dist} \tan(\beta) = -\tan(\beta') \quad (43)$$

Lorsque l'onde est considérée comme plane, deux arêtes parallèles d'un bâtiment apparaissent également parallèles sur l'image RSO. De la même façon, il est possible de montrer que $p_{AC}^{RADAR} = p_{BD}^{RADAR}$. Le rectangle ABCD en géométrie terrain devient un parallélogramme en géométrie RADAR.

Si l'onde n'est pas plane, les incidences locales à chaque extrémité du quadrilatère doivent être considérées. Soit X_x la distance D_x projetée au sol à l'incidence θ_x .

En géométrie terrain on a :

$$X_B = X_A - L \cos(\beta)$$

$$X_C = X_A + l \sin(\beta)$$

$$X_D = X_A + l \sin(\beta) - L \cos(\beta)$$

et

$$Y_B = Y_A - L \sin(\beta)$$

$$Y_C = Y_A - l \cos(\beta)$$

$$Y_D = Y_A - l \cos(\beta) - L \sin(\beta)$$

Les angles d'incidence en B , C et D s'expriment en fonction de l'angle d'incidence en A (θ_A) et des caractéristiques du bâtiment :

$$\theta_B = \arctan\left(\tan(\theta_A) - \frac{L}{H} \cos(\beta)\right)$$

$$\theta_C = \arctan\left(\tan(\theta_A) + \frac{l}{H} \sin(\beta)\right)$$

$$\theta_D = \arctan\left(\tan(\theta_A) + \frac{l \sin(\beta) - L \cos(\beta)}{H}\right)$$

Les pentes des segments en géométrie RADAR sans hypothèse d'onde plane sont :

$$p_{AB}^{RADAR} = \frac{(Y_B - Y_A)}{(D_B - D_A)} \frac{pas_azi}{pas_dist}$$

$$p_{AB}^{RADAR} = \frac{L \sin(\beta)}{H \left(\frac{1}{\cos(\theta_B)} - \frac{1}{\cos(\theta_A)} \right)} \frac{pas_azi}{pas_dist} = -\tan(\beta'_{AB}) \quad (44)$$

et

$$p_{CD}^{RADAR} = \frac{L \sin(\beta)}{H \left(\frac{1}{\cos(\theta_D)} - \frac{1}{\cos(\theta_C)} \right)} \frac{pas_azi}{pas_dist} = -\tan(\beta'_{CD}) \quad (45)$$

Considérons une acquisition telle que $H = 3000\text{m}$ et $D_A = 4000\text{m}$. Les pentes en géométrie RADAR des segments AB et CD sont calculées pour des bâtiments de 10 à 80m. Sur la figure 14 est représenté la différence entre β'_{AB} et β'_{CD} en fonction de β lorsque les pas d'échantillonnage en distance et azimut sont égaux.

Lorsque β est proche de 0° ou 90° , $\beta'_{AB} \approx \beta'_{CD}$. Lorsque l augmente, il est visible que la différence d'orientation augmente également. Cette évolution est normale car lorsque l augmente, la différence entre θ_A et θ_C ainsi que celle entre θ_B et θ_D augmente. Néanmoins, même pour de grands bâtiments ($L=l=80\text{m}$), cette différence reste inférieure au degré.

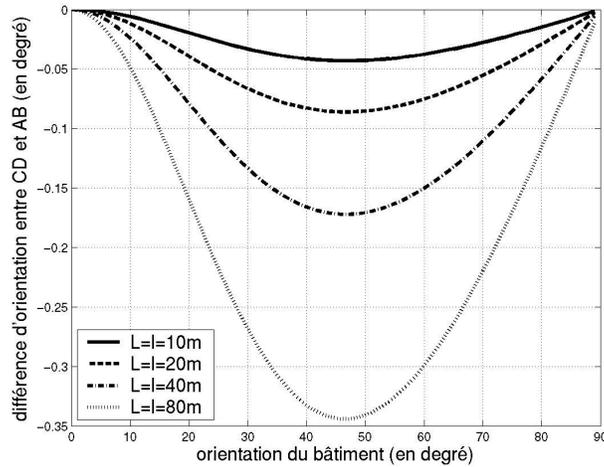


Figure 14 – parallélisme des segments AB et CD pour un rapport de pas d'échantillonnage de 1

Pour des pas d'échantillonnage très différents (0.5 et 1.5), l'évolution du parallélisme des segments AB et CD est représenté sur la figure 15. La valeur de β pour laquelle $\beta'_{AB} - \beta'_{CD}$ est maximal varie en fonction des pas d'échantillonnage. La différence d'orientation entre les deux segments reste toujours inférieure au degré et les segments peuvent être considéré parallèles.

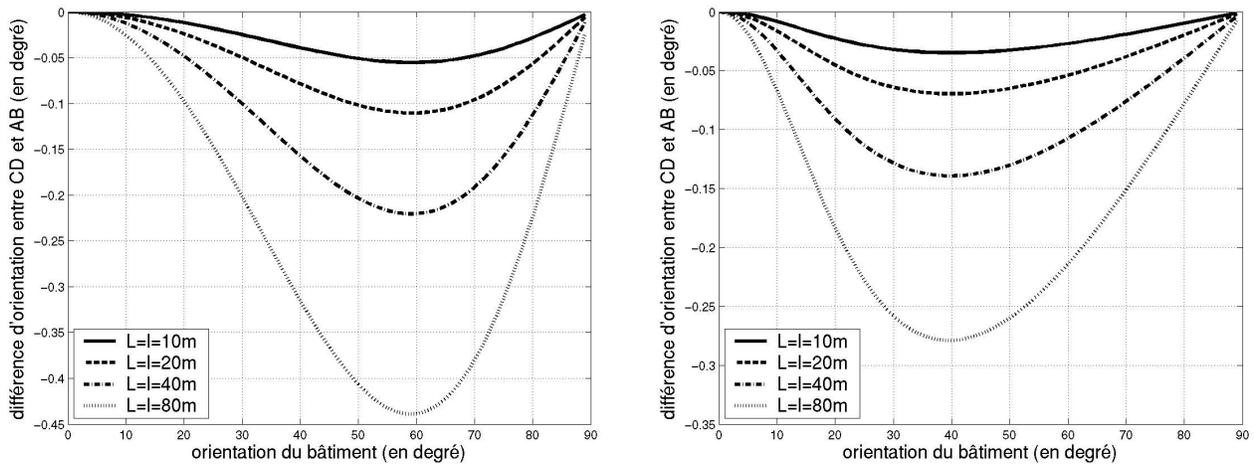


Figure 15 – parallélisme des segments AB et CD pour un rapport de pas d'échantillonnage de a) 0.5 et b) 1.5

2.2.3. Distorsions géométriques

L'acquisition en distance, effectuée par le capteur avec un angle de visée non nul, a des conséquences sur la représentation des objets élevés (figure 16).

La compression apparaît lorsque le relief de l'objet présente une pente face au capteur et inférieure à l'incidence de l'onde RADAR. Le segment A_2B_2 apparaît plus petit sur l'image que A_1B_1 ne l'est en réalité.

Le repliement apparaît lorsque le relief de l'objet présente une pente face au capteur et supérieure à l'incidence de l'onde RADAR. Le point D_1 , d'altitude plus élevée est alors plus proche du capteur que C_1 . Par conséquent, l'ordonnancement des points de cette face est inversé par rapport à la réalité. Ainsi D_2 apparaît sur l'image avant C_2 . Les effets du

repliement ne se limitent pas à la géométrie et sont étudiés en détail au §2.5 qui sera dédié aux effets radiométriques.

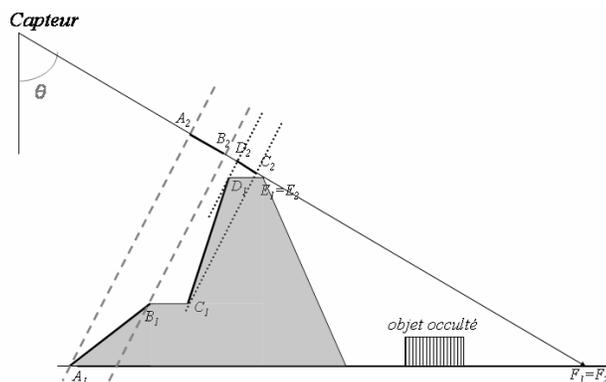


Figure 16 – Distorsions géométriques engendrées par l'acquisition en distance

2.2.4. Les ombres

L'ombre apparaît lorsque le relief de l'objet présente une pente opposée au capteur avec un angle supérieur à l'angle d'incidence de l'onde RADAR. Cette partie de l'objet ne renvoie pas d'informations au capteur. Les points E_1 et F_1 (cf. figure 16) sont alors vus sous le même angle par le capteur mais ne sont pas à la même distance. Entre ces deux points, aucune onde RADAR n'est reçue par le capteur. Pendant ce laps de temps, les données reçues correspondent à du bruit. Des phénomènes d'occultations apparaissent pour les objets qui se situent dans la zone d'ombre et qui ne sont pas vus par le capteur. Les ombres sont présentes sur les images en amplitude, l'interférogramme et la cohérence.

2.2.5. Conclusion

Une image RADAR peut se visualiser de deux façons : en géométrie RADAR et en géométrie terrain. La géométrie RADAR présente l'avantage d'avoir un pas d'échantillonnage constant. En géométrie terrain la taille du pixel dépend de l'angle d'incidence et l'image nécessite alors un ré-échantillonnage pour obtenir une taille de pixel constante. L'hypothèse d'onde localement plane est vérifiée ce qui facilite les représentations et les calculs. La géométrie d'acquisition en distance provoque d'importantes distorsions qui rendent l'interprétation de la scène difficile.

2.3. Mesures du 3D en imagerie RSO interférométrique

Dans ce paragraphe sont présentées les deux mesures (*monoscopiques et interférométriques*) permettant d'extraire une information 3D à partir d'un couple interférométrique. Les précisions altimétriques et planimétriques de ces mesures sont comparées.

2.3.1. Les mesures monoscopiques

Les mesures monoscopiques sont obtenues en utilisant une seule image. Elles sont effectuées sur les ombres et sur les façades des bâtiments. L'extraction du 3D par monoscopie est une mesure relative de hauteur. Pour en connaître une mesure absolue, il faut déterminer le modèle numérique de terrain et ainsi obtenir l'altitude du sol au niveau du bâtiment.

2.3.1.1. Mesure monoscopique sur les ombres

L'analyse de la forme des ombres peut apporter des informations sur les murs du bâtiment [Bolter 2001], à savoir :

- la forme du toit.
- la hauteur h_B de la partie du bâtiment projetant l'ombre (dans le cas présent le sol est supposé plat).

h_B est obtenue par mesure monoscopique en géométrie RADAR grâce à la longueur s de l'ombre selon l'axe des distances (cf. figure 17) par :

$$h_B = s \cdot \cos(\theta) \quad (46)$$

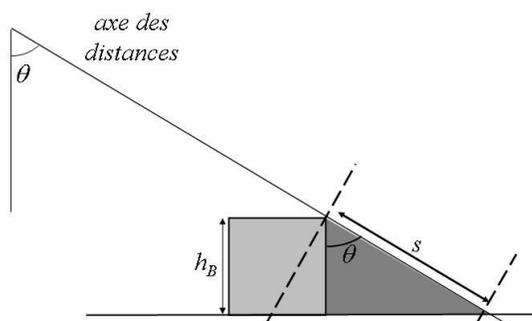


Figure 17 – Mesure monoscopique sur les ombres

L'expression de la précision sur la hauteur des bâtiments pour une erreur de mesure de δn pixels sur un coté de l'ombre est donnée par :

$$\delta h_B = \delta s \cdot \cos(\theta) = res_{dist} \cdot \delta n \cdot \cos(\theta) \quad (47)$$

L'erreur de localisation planimétrique du toit pour une erreur de détection δm sur le contour entre le bord du bâtiment et l'ombre est :

$$\delta x = res_{dist} \cdot \delta m \cdot \sin(\theta) \quad (48)$$

L'angle d'incidence influe sur la précision de la mesure. Plus il est grand, plus la précision altimétrique sur la mesure monoscopique des ombres (47) est bonne. La précision planimétrique a une évolution contraire et diminue lorsque l'angle d'incidence augmente (48). Pour un angle d'incidence de 45° , les erreurs planimétriques et altimétriques sont les mêmes.

2.3.1.2. Mesure de la parallaxe

La mesure de parallaxe permet de déterminer la hauteur de façades verticales de bâtiments faisant face au capteur. Cette mesure nécessite des informations sur le bas et le haut de la façade (cf. figure 18).

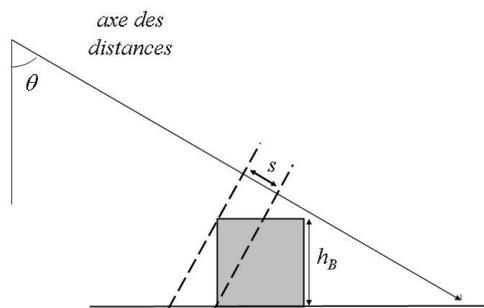


Figure 18 – Mesure monoscopique sur une façade

La hauteur h_B du bâtiment par mesure monoscopique de parallaxe est donnée par :

$$h_B = \frac{s}{\cos(\theta)} \quad (49)$$

où s est la mesure du repliement selon l'axe des distances.

La précision altimétrique sur cette mesure pour une erreur de pointé de δn pixels sur le début ou la fin de la façade est donnée par :

$$\delta h_B = \frac{\delta s}{\cos(\theta)} = \frac{res_{dist} \cdot \delta n}{\cos(\theta)} \quad (50)$$

Le bas de la façade donne la position au sol du mur le plus proche du capteur. Pour une erreur de mesure de δm pixels, l'erreur planimétrique est obtenue par :

$$\delta x = res_{dist} \cdot \delta m \sin(\theta) \quad (51)$$

L'évolution de la précision altimétrique de la mesure de parallaxe en fonction de l'angle d'incidence est opposée à celle de la mesure monoscopique sur les ombres. Plus l'angle d'incidence sera faible, meilleure sera la précision altimétrique sur la mesure de parallaxe (50). La précision planimétrique de la mesure de la parallaxe est identique à celle de la mesure monoscopique sur les ombres.

2.3.1.3. Limitations

Les mesures monoscopiques ne sont pas toujours valides. Dans les deux exemples présentés sur la figure 19, la mesure monoscopique sur les ombres n'est pas utilisable (la relation (46) n'est plus vérifiée) à cause du repliement. Dans le cas de la figure 19a le bas de la façade coupe une partie de l'ombre. Sur la figure 19b le bâtiment 2 masque une partie de l'ombre. Dans les deux cas, la hauteur du bâtiment fournie par cette mesure sous estime la hauteur réelle du bâtiment.

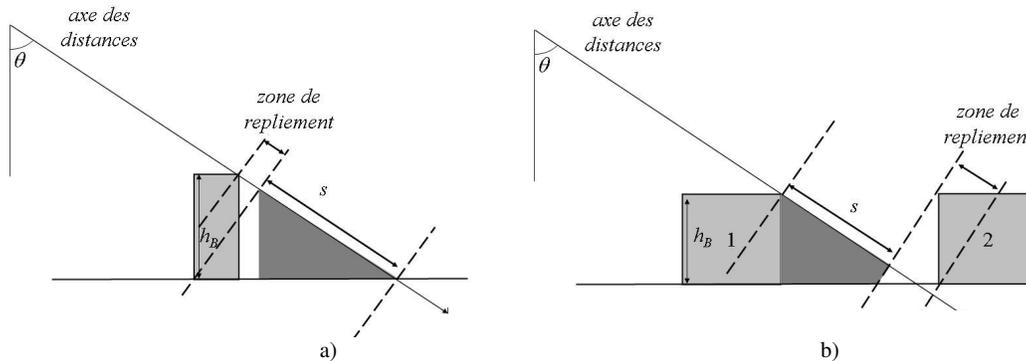


Figure 19 – Limitation de la mesure monoscopique sur les ombres

La mesure monoscopique sur la parallaxe ne peut pas être utilisée dans le cas présenté sur la figure 20. En effet, le haut du toit à double pente apparaissant au début de la zone de repliement n'appartient pas à la façade verticale du bâtiment.

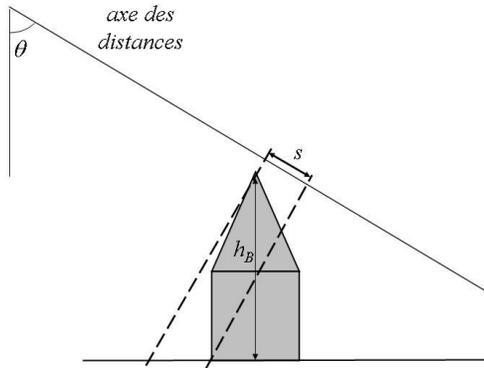


Figure 20 – Limitation de la mesure monoscopique sur la parallaxe

2.3.2. Comparaisons des précisions des mesures monoscopiques et interférométriques

En imagerie RSO aéroporté, l'angle d'incidence local varie beaucoup le long de la fauchée. Il est calculé par :

$$\theta(D) = \cos^{-1}\left(\frac{H}{D}\right) \quad (52)$$

avec : D distance capteur / cible
 H hauteur du capteur

Les précisions altimétriques et planimétriques sur les mesures monoscopiques des images RAMSES, présentées en annexe A, sont détaillées sur la figure 21 le long de la fauchée.

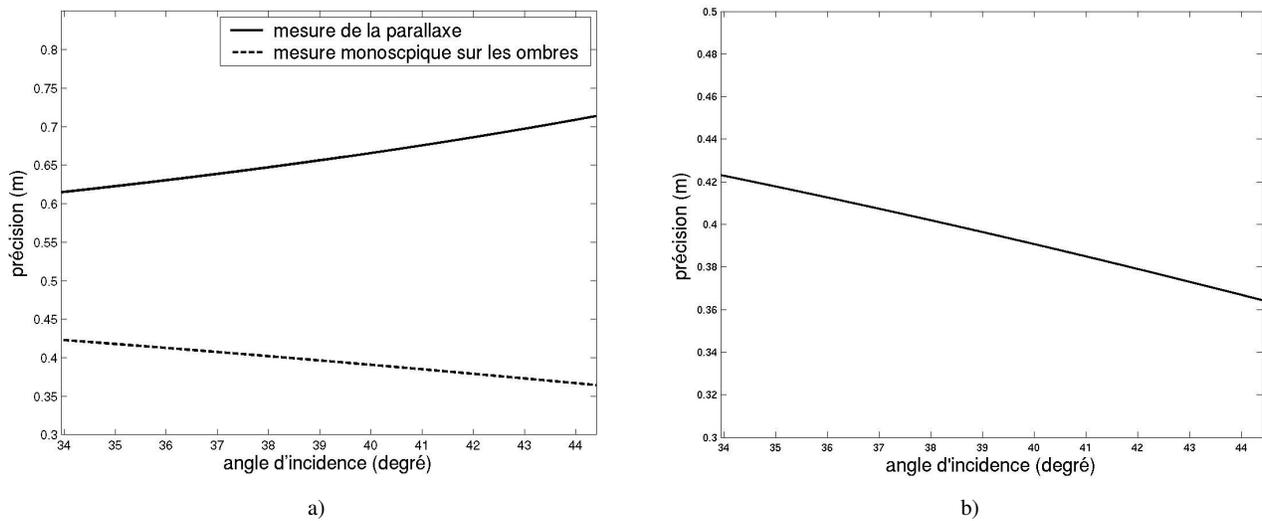


Figure 21 – Précisions sur les mesures monoscopiques
 a) altimétriques (formules 47 et 50)
 b) planimétriques (formules 48 ou 51)

Les précisions altimétriques sur les réflexions spéculaires et le chatoiement de Rayleigh sont présentées sur la figure 22 pour différents RSB (cf. §1.3.5). Dans les deux cas, les bruits surfaciques et thermiques sont pris en compte. Le décalage fréquentiel provenant du bruit surfacique est très faible par rapport à la largeur de bande sur les images RAMSES haute résolution. Par conséquent, ce terme sera négligé. La hauteur d'ambiguïté variant entre le

début et la fin de la fauchée, les précisions altimétriques entre le début et la fin de la fauchée sont représentées.

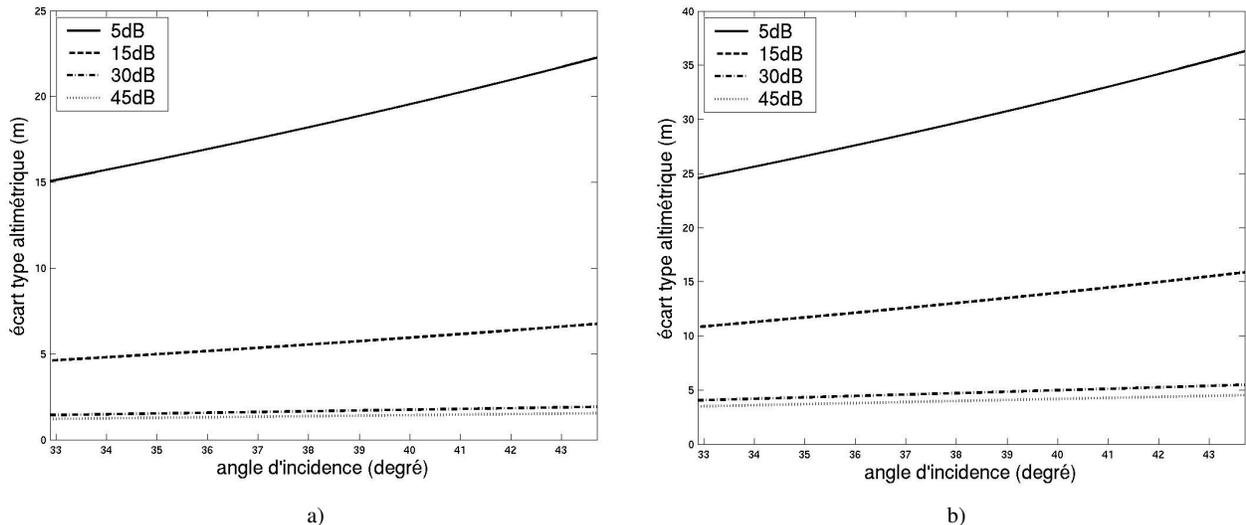


Figure 22 – Précisions altimétriques sur la mesure interférométrique pour a) une réflexion spéculaire (formule 39) et b) une rétrodiffusion de type Rayleigh (formule 38) de RSB 5, 15, 30 et 45dB

La précision altimétrique de la mesure monoscopique sur les ombres est très bonne et inférieure à la taille de la cellule de résolution des images. La précision de la mesure sur la parallaxe est plus faible mais reste bonne. Les précisions planimétriques des deux mesures sont, elles aussi, inférieures à la taille de la cellule de résolution du capteur.

Les mesures interférométriques sont les moins précises même pour de forts RSB. Pour des réflexions spéculaires, l'écart-type est de 1.5m environ lorsque le RSB est supérieur à 30dB. Pour les rétrodiffuseurs de type Rayleigh, il est toujours supérieur à 4m.

L'utilisation du multivue permet d'améliorer ces précisions d'un facteur \sqrt{L} où L est le nombre de vue. Les tests effectués sur le site de Dunkerque montrent que le niveau de RSB du sol est environ de 10dB. Nous avons tracé ci-dessous la précision altimétrique pour un tel RSB en fonction du nombre de vues pour une rétrodiffusion de type Rayleigh. Pour un nombre élevé de vue ($L=36$), la précision sur l'altitude du sol est de 4m seulement. Néanmoins un moyennage sur une grande zone permet d'obtenir de meilleures précisions.

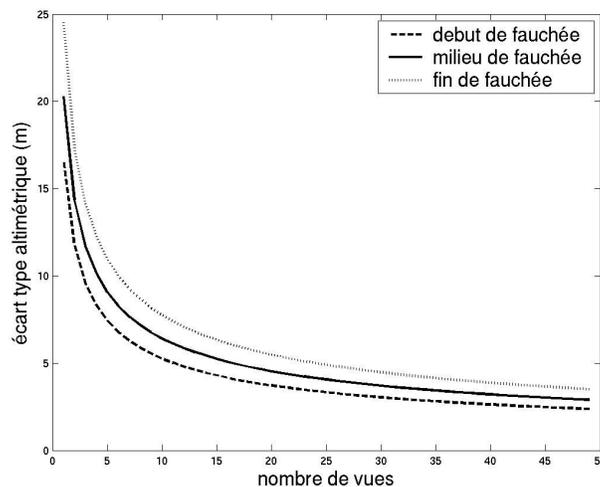


Figure 23 – Influence du nombre de vues sur la précision interférométrique pour une rétrodiffusion de type Rayleigh

2.3.3. Conclusion

Deux types de mesures du 3D à partir de données interférométriques ont été présentées : les mesures monoscopiques et la mesure interférométrique. La mesure monoscopique sur les ombres offre la meilleure précision altimétrique. La mesure de la parallaxe est légèrement moins précise que la mesure monoscopique sur les ombres mais l'est beaucoup plus que la mesure interférométrique. Les précisions altimétriques et planimétriques des mesures monoscopiques sont de l'ordre de la taille de la cellule de résolution. La précision altimétrique des mesures interférométriques est de quelques mètres dans les conditions d'acquisition des images RAMSES présentées en annexe A. Néanmoins, un moyennage spatial sur de grandes zones permet d'améliorer cette précision. Les mesures monoscopiques sur les ombres peuvent, dans certains cas, ne pas être valides à cause du repliement. Pour que la mesure monoscopique sur la parallaxe soit utilisable, il faut que les deux primitives considérées appartiennent à la même façade.

Les avantages et inconvénients de chaque mesure peuvent être résumés dans le tableau suivant :

	avantages	inconvénients
mesure monoscopique sur les ombres	très bonne précision	ombres pas toujours visibles ou tronquées
mesure monoscopique sur les échos	bonne précision	pas toujours d'échos du bord du toit franc
interférométrie	mesure dense	mauvaise précision

2.4. Analyse phénoménologique

2.4.1. Introduction

Les caractéristiques électromagnétiques de l'onde RADAR influent sur les phénomènes apparaissant sur les images. Selon la longueur d'onde utilisée et la nature de la surface éclairée (rugosité, composition, orientation...), l'onde ne se comporte pas de la même façon. De plus, en milieu urbain et en bande X, l'onde électromagnétique est suffisamment énergétique pour subir de multiples rebonds, complexifiant ainsi la compréhension de la scène.

2.4.2. L'équation RADAR

L'équation RADAR permet de relier la puissance émise et la puissance reçue par le capteur après avoir atteint une cible. L'émetteur et le récepteur constituent le même dispositif donc les gains d'émission et de réception sont supposés identiques et sont notés G.

Dans le cas de cibles ponctuelles, l'équation RADAR s'écrit :

$$P_R = P_E \frac{G^2 \lambda^2 \sigma}{(4\pi)^3 D^4}$$

avec P_R, P_E puissance reçue et émise
 D distance du capteur à la cible
 λ longueur d'onde du signal RADAR
 G gain de l'antenne
 σ surface équivalente RADAR (SER)

Cette équation se décompose en trois termes :

- $P_E \frac{G}{4\pi D^2}$ traduit la puissance à la distance D émise par une antenne directionnelle de gain G.

- $\frac{\sigma}{4\pi D^2}$ décrit la puissance de l'onde rétrodiffusée par une cible à la distance D
- $\frac{G\lambda^2}{4\pi}$ représente le facteur d'énergie capté par l'antenne.

La SER représente la portion d'énergie renvoyée par une cible ponctuelle. Elle s'exprime en m^2 . Dans le cas de cibles étendues, le coefficient de rétrodiffusion, noté σ^0 , est défini comme la valeur moyenne de la SER par unité de surface. Il s'exprime en m^2/m^2 ou, en échelle logarithmique, en dBm^2/m^2 .

Le σ^0 dépend de la composition de la cible, de l'angle d'incidence et de la longueur d'onde.

2.4.3. Notion de rugosité et de textures

La rugosité d'une surface en imagerie RSO se définit par rapport à la longueur d'onde (λ) et l'angle d'incidence (θ) de l'onde électromagnétique. Néanmoins, la limite entre une zone rugueuse et une zone lisse reste subjective. Un critère couramment utilisé est le critère de Rayleigh. Il tient compte de la hauteur des irrégularités des surfaces mais pas de leur densité.

Soit e la hauteur de ces irrégularités. Une surface est rugueuse d'après le critère de Rayleigh si :

$$e > \frac{\lambda}{8\cos(\theta)}$$

Plus la longueur d'onde utilisée est faible, plus les petites irrégularités sont détectées et plus l'image RSO en amplitude est texturée. En bande X ($\lambda=3.17cm$) pour une incidence de 45° , la surface est rugueuse si les déformations font plus de 0.56cm d'épaisseur.

Une surface rugueuse se traduit sur l'image en amplitude, par des textures assimilables, en première approximation, à du chatolement pleinement développé (cf. annexe B.2.1). Dans ce cas, l'onde électromagnétique subit une rétrodiffusion de Rayleigh. Le niveau de rétrodiffusion peut-être plus ou moins important selon la nature de la surface rétrodiffusante.

Sur la figure 24a, le toit du bâtiment, apparaît avec un niveau de rétrodiffusion plus élevé que le sol.

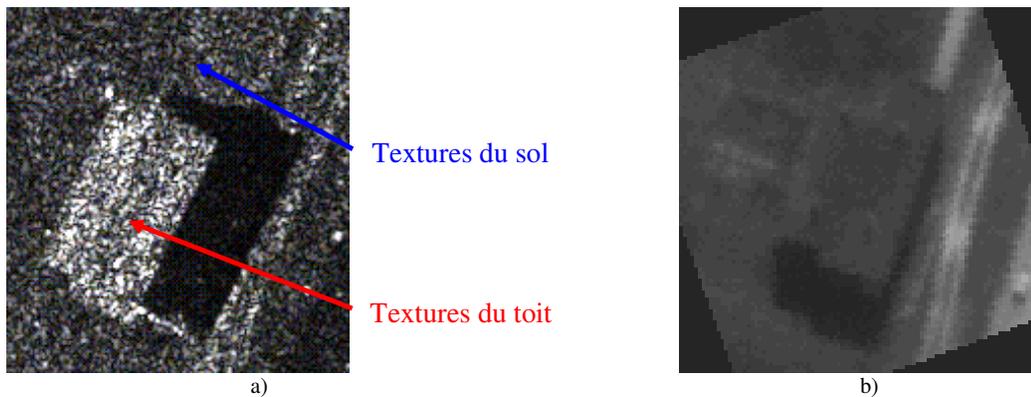


Figure 24 –a) Exemple de textures, b) photo aérienne du même bâtiment dans le domaine du visible

Sur l'image optique correspondante, figure 24b, le bâtiment ne présente pas de textures et, est peu contrasté par rapport au sol. En revanche, sur l'image RSO, le toit présente des

rugosités suffisantes pour engendrer des textures très contrastées. Détecter ce bâtiment semble donc plus difficile sur cette image optique que sur l'image RSO à disposition.

2.4.4. Les échos

2.4.4.1. Les échos simples

Lorsque les irrégularités de la structure éclairée par l'onde électromagnétique sont inférieures à la limite fixée par le critère de Rayleigh, on ne parle plus de rétrodiffusion mais de réflexion spéculaire.

Les échos forts simples apparaissent lorsque l'onde électromagnétique effectue une réflexion spéculaire en direction du capteur. Par exemple des pièces métalliques (climatisations, gouttières, pipelines, pylônes électriques etc...) peuvent être à l'origine de ces échos. Le coefficient de rétrodiffusion de ce type de cible est positif et peut atteindre 50dBm²/m². Les échos simples apparaissent dans la case distance où a lieu la réflexion. Ils ont une amplitude qui dépend de l'angle de réflexion de l'onde RADAR. Lorsqu'il n'y a pas de repliement, les échos simples ont l'altitude de la cellule de résolution où a eu lieu la réflexion.

Les réflexions sont supposées suivre la loi de Descartes. Par conséquent, si la surface lisse est orientée vers le capteur (figure 25a), l'écho apparaît avec une forte amplitude. Plus le lobe d'antenne du signal réfléchi est dirigé vers le capteur, plus l'amplitude de l'écho sera élevée. Si l'onde subit une réflexion sur une surface horizontale (figure 25b), elle peut ne pas revenir vers le capteur.



Figure 25 – Représentation de la direction principale (flèche) et du lobe d'antenne de l'onde RADAR après avoir subi une réflexion simple sur une surface lisse a) orientée vers le capteur et b) horizontale

Sur la figure 26 sont présentés des exemples où des échos simples sont visibles.

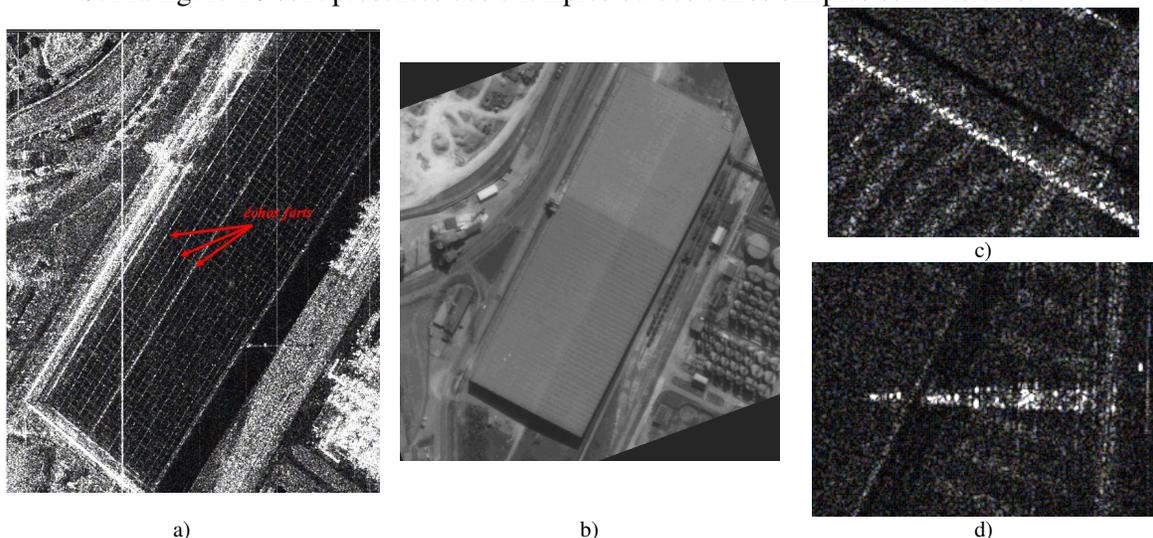


Figure 26 – a) Exemples d'échos simples sur le toit d'un bâtiment sur image en amplitude RSO, b) même bâtiment sur image optique, c) pipeline et d) pylône sur l'image en amplitude RSO

Sur la figure 26a, selon les deux axes du bâtiment, des jonctions régulières, entre les plaques de métal qui constituent le toit, sont visibles. L'amplitude des échos résultant sur le toit est relativement faible car la surface réfléchissante n'est pas orientée vers le capteur. Sur l'image optique correspondante (figure 26b), seules les jonctions suivant la largeur du bâtiment sont visibles. Les autres jonctions ne sont pas visibles sans doute à cause de la direction de l'éclairage.

Sur les images c et d des exemples d'échos simples, réfléchis par un pipeline et un pylône, sont présentés.

2.4.4.2. Les échos doubles

Les échos doubles subissent deux réflexions / rétrodiffusions et apparaissent lorsque la scène comporte des dièdres ou des coins réflecteurs orientés vers le capteur. En zone urbaine, un dièdre est formé entre la façade d'un bâtiment parallèle à l'axe des azimuts et le sol. La figure 27 représente une vue du dessus d'une scène avec deux bâtiments. La façade rouge du bâtiment et le sol constituent un dièdre.

Lorsque le bâtiment n'est pas parallèle à l'azimut, le terme de pseudo-dièdre sera employé. La façade bleue du bâtiment B et le sol est un exemple de pseudo-dièdre. Dans [Simonetto 2002], un modèle pour les échos doubles a été proposé. Dans ce paragraphe, nous nous proposons d'utiliser ce modèle pour déterminer la hauteur interférométrique des échos doubles.

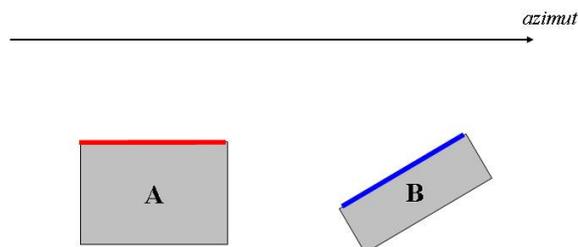


Figure 27 – Dièdre et pseudo-dièdre

2.4.4.2.1. Modèle dans le cas diédrique

Les échos doubles diédriques sont modélisés par une somme d'échos ayant subi une réflexion miroir puis une rétrodiffusion vers le capteur. Ces réflexions/rétrodiffusions peuvent avoir lieu sur l'intégralité de la façade du bâtiment. Deux exemples de ces trajets sont représentés sur la figure 28 lorsque le bâtiment est parallèle à la trajectoire du capteur. Lorsque l'onde est supposée plane, c'est-à-dire le capteur supposé à l'infini, l'onde subit deux réflexions miroirs sur le sol et le mur. Peu importe la hauteur h où a lieu la réflexion sur le mur et l'angle d'incidence, le trajet parcouru est équivalent au trajet simple capteur / dièdre.

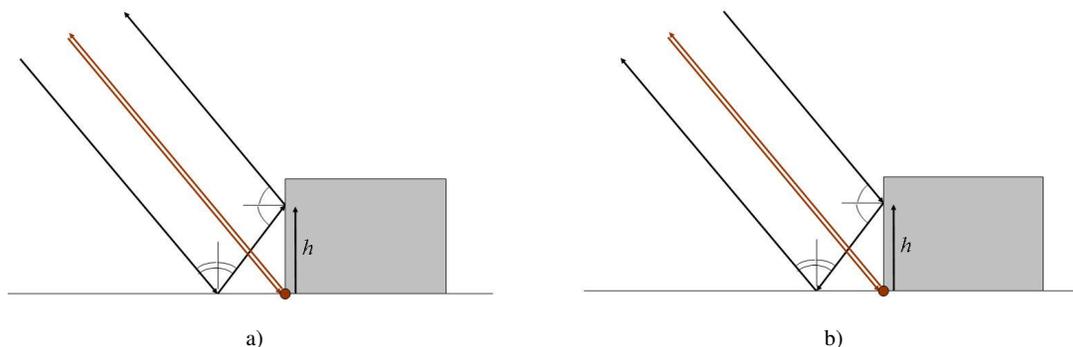


Figure 28 – Trajets effectués par les échos diédriques avec deux réflexions miroir

Si le capteur n'est pas à l'infini (onde sphérique), après deux réflexions, l'onde n'est pas reçue par le capteur. En revanche si l'onde subit une réflexion et une rétrodiffusion, elle peut revenir vers le capteur (cf. figure 29). Dans le cas a, la réflexion a lieu sur le sol et est suivie d'une rétrodiffusion sur le mur. Dans le cas b, la réflexion miroir a lieu sur le mur et est suivie d'une rétrodiffusion sur le sol.

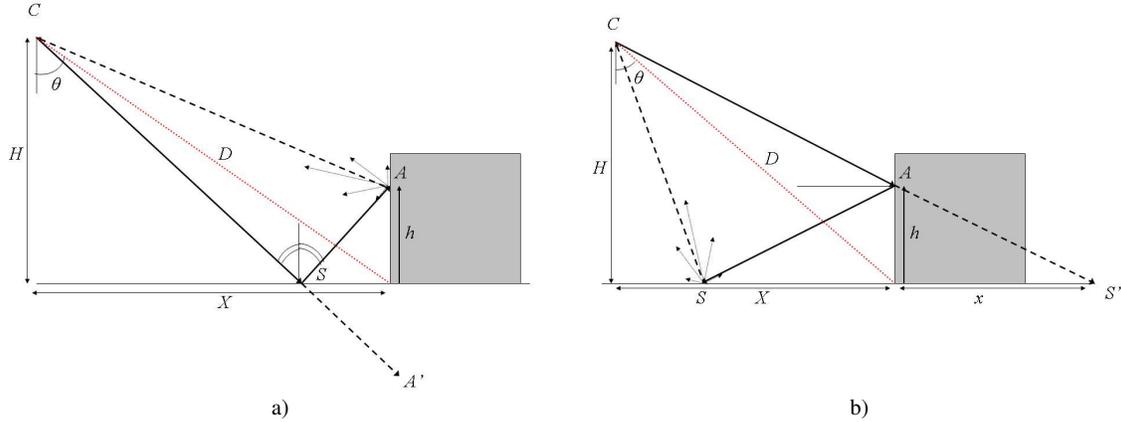


Figure 29 – Trajets effectués par les échos diédriques avec réflexion miroir a) sur le mur, b) sur le sol

Le calcul de la différence de trajet effectué par l'onde lorsque le capteur est ou n'est pas à l'infini est proposé dans [Simonetto 2002] en géométrie RADAR. Pour cela, la différence de parcours, notée Δ , entre le trajet simple capteur-dièdre et le trajet double est calculée. Nous proposons de calculer ce même trajet en géométrie terrain. L'intérêt de cette approche est de considérer un unique angle θ correspondant à l'angle d'incidence de l'onde au niveau du dièdre.

cas a : le trajet parcouru par l'onde après une réflexion miroir sur le sol est (cf. figure 29a) :

$$D_{double} = CA + AS + SC$$

En considérant le symétrique de A par rapport au sol on a :

$$D_{double} = CA + A'C$$

Avec :

$$CA = \sqrt{X^2 + (H - h)^2}, A'C = \sqrt{X^2 + (H + h)^2}$$

$$D = \sqrt{X^2 + H^2}$$

$$\Delta(h) = D_{double} - D$$

$$\Delta(h) = \sqrt{X^2 + H^2} - \frac{1}{2} \left(\sqrt{X^2 + (H - h)^2} + \sqrt{X^2 + (H + h)^2} \right)$$

$$\Delta(h) = \sqrt{X^2 + H^2} \left(1 - \frac{1}{2} \sqrt{1 + \frac{h^2 + 2hH}{X^2 + H^2}} - \frac{1}{2} \sqrt{1 + \frac{h^2 - 2hH}{X^2 + H^2}} \right) \quad (53)$$

En aéroporté, H et X sont très grands devant h . On peut effectuer un développement limité à l'ordre 2 de l'expression (53).

$$\Delta(h) \approx \frac{h^2 X^2}{2(X^2 + H^2)^{3/2}} \quad (54)$$

De plus, l'angle d'incidence au niveau du dièdre s'exprime par :

$$\sin(\theta) = \frac{X}{\sqrt{(X^2 + H^2)}}$$

D'où :

$$\Delta(h) \approx \frac{h^2 \sin^2(\theta)}{2D} \quad (55)$$

cas b : le trajet double parcouru par l'onde est (cf. figure 29b) :

$$D_{double} = CA + AS + SC$$

En considérant le symétrique de S par rapport au mur on a :

$$D_{double} = CS' + SC$$

De plus on a :

$$CS' = \sqrt{(X+x)^2 + H^2}, SC = \sqrt{(X-x)^2 + H^2}$$

$$D = \sqrt{X^2 + H^2}$$

avec :

$$x = X \frac{h}{H-h}$$

D'où :

$$\Delta(x) = \sqrt{X^2 + H^2} - \frac{1}{2} \left(\sqrt{(X+x)^2 + H^2} + \sqrt{(X-x)^2 + H^2} \right)$$

$$\Delta(h) = \sqrt{X^2 + H^2} - \frac{1}{2} \left(\sqrt{X^2 \left(1 + \frac{h}{H-h}\right)^2 + H^2} + \sqrt{X^2 \left(1 - \frac{h}{H-h}\right)^2 + H^2} \right)$$

$$\Delta(h) = \sqrt{X^2 + H^2} - \frac{1}{2} \left(\sqrt{X^2 \left(\frac{H}{H-h}\right)^2 + H^2} + \sqrt{X^2 \left(\frac{H-2h}{H-h}\right)^2 + H^2} \right) \quad (56)$$

Un développement limité à l'ordre 2 de l'expression (56) donne

$$\Delta(h) = h^2 \frac{\sin^2(\theta)}{2D} \quad (57)$$

où D est la distance entre le capteur et le dièdre et θ l'angle d'incidence local de l'onde RADAR au niveau du dièdre

Pour les deux trajets, pour une rétrodiffusion (cas a) ou une réflexion (cas b) à la hauteur h du bâtiment, la différence de trajet parcouru entre l'écho ayant effectué deux rebonds et l'écho simple capteur / dièdre est le même aux deuxième ordre. Néanmoins ces deux trajets ne sont pas strictement identiques : un développement limité à l'ordre 3 de la différence (53) et (56) donne une différence de parcours $\Delta'(h)$:

$$\Delta'(h) = h^3 \frac{\tan(\theta) \sin(\theta)}{D^2} \quad (58)$$

Il est intéressant de remarquer que les relations (42) p33 et (57) sont similaires car la même hypothèse est effectuée sur θ . Sur la figure 30 sont représentées les localisations des échos doubles avec hypothèse d'onde plane (θ constant) et sans cette hypothèse.

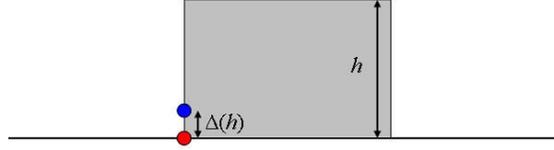


Figure 30 – localisation de l'écho double ayant atteint le mur à la hauteur h avec hypothèse d'onde plane (rouge) et sans hypothèse d'onde plane (bleu)

Nous nous proposons d'utiliser ce modèle en interférométrie pour déterminer l'altitude de l'écho diédrique. Sur la figure 31 est représentée la configuration d'acquisition interférométrique pour deux points P et Q appartenant à une façade. La hauteur du bâtiment est notée h_B .

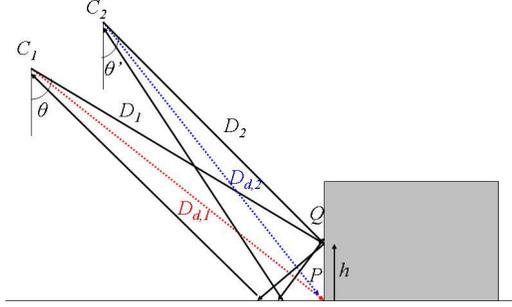


Figure 31 – Schéma de prise de vue d'un interférogramme sur une façade

Les composantes de l'écho diédrique ayant subi une réflexion/rétrodiffusion à la hauteur h du bâtiment s'expriment par :

$$A_1(h) \exp\left(-j\left(\frac{2\pi}{\lambda}[D_{d,1} + \Delta_1(h)] + \varphi_1(h)\right)\right) \text{ et } A_2(h) \exp\left(-j\left(\frac{2\pi}{\lambda}[D_{d,2} + \Delta_2(h)] + \varphi_2(h)\right)\right)$$

où :

$$\Delta_1(h) = h^2 \frac{\sin^2(\theta)}{2D_1} \text{ et } \Delta_2(h) = h^2 \frac{\sin^2(\theta')}{2D_2}$$

et φ_i est un terme provenant du trajet de l'onde entre les diffuseurs à l'intérieur de la cellule de résolution où il y a une rétrodiffusion. Ce terme est inconnu et supposé aléatoire mais reproductible. Nous supposons que $\varphi_1(h) = \varphi_2(h)$ car, pour une même hauteur, l'onde a atteint les mêmes cellules de résolution sur le sol et la façade.

Les contributions sur la façade de hauteur h_B , pour chaque image du couple interférométrique, peuvent s'écrire :

$$\int_0^{h_B} A_1(h) \exp\left(-j\left(\frac{2\pi}{\lambda}[D_{d,1} + \Delta_1(h)] + \varphi_1(h)\right)\right) dh \text{ et } \int_0^{h_B} A_2(h) \exp\left(-j\left(\frac{2\pi}{\lambda}[D_{d,2} + \Delta_2(h)] + \varphi_2(h)\right)\right) dh$$

En formant l'interférogramme on obtient l'expression du vecteur interférométrique d'un écho diédrique :

$$\exp\left(-j\frac{2\pi}{\lambda}(D_{d,1} - D_{d,2})\right) \left[\int_0^{h_B} A_1(h) \exp\left(-j\left(\frac{2\pi}{\lambda}\Delta_1(h) + \varphi_1(h)\right)\right) dh \right] \left[\int_0^{h_B} A_2(h) \exp\left(j\left(\frac{2\pi}{\lambda}\Delta_2(h) + \varphi_2(h)\right)\right) dh \right]$$

Le premier terme de l'équation précédente correspond à l'altitude du dièdre s'il n'y avait qu'un trajet sans doubles réflexions. Le deuxième terme, un produit de deux intégrales

complexes, correspond à la contribution interférométrique provenant de l'ensemble des doubles rebonds. A cause des composantes aléatoires, ces intégrales ne peuvent pas être calculées analytiquement. Elles sont donc discrétisées en choisissant un pas d'échantillonnage, noté pas , suffisamment faible pour que le théorème de Shannon soit respecté par rapport à λ . Nous supposons que l'amplitude des contributions $A_1(h)$ et $A_2(h)$ sont identiques le long de la façade et valent A . Le produit des deux intégrales s'exprime par :

$$A^2 \left[\sum_{k_1=0}^{n_B} \exp \left(-j \left(\frac{2\pi}{\lambda} \Delta_1(k_1 pas) + \varphi_1(k_1 pas) \right) \right) \right] \left[\sum_{k_2=0}^{n_B} \exp \left(j \left(\frac{2\pi}{\lambda} \Delta_2(k_2 pas) + \varphi_2(k_2 pas) \right) \right) \right]$$

où $h_B = pas.n_B$.

En isolant les termes avec une phase aléatoire identique, le produit des deux intégrales devient :

$$A^2 \left[\sum_{k_1=0}^{n_B} \exp \left(-j \frac{2\pi}{\lambda} (\Delta_1(k_1 pas) - \Delta_2(k_1 pas)) \right) \right] + A^2 \left[\sum_{k_1=0}^{n_B} \sum_{\substack{k_2=0 \\ k_2 \neq k_1}}^{n_B} \exp \left(-j \left(\frac{2\pi}{\lambda} (\Delta_1(k_1 pas) - \Delta_2(k_2 pas)) + \varphi_1(k_1 pas) - \varphi_2(k_2 pas) \right) \right) \right]$$

L'expression à l'intérieur de l'exponentielle du terme de droite est aléatoire à cause des termes en φ_1 et φ_2 . En moyenne, le terme de droite est donc nul. La hauteur de l'écho double par rapport au dièdre s'exprime par :

$$\arg \left[\sum_{k_1=0}^{n_B} \exp \left(-j \frac{2\pi}{\lambda} (\Delta_1(k_1 pas) - \Delta_2(k_1 pas)) \right) \right]$$

où \arg correspond à l'argument du nombre complexe.

Pour un porteur volant à 3000m d'altitude, avec deux antennes formant une base verticale de 50cm, une longueur d'onde de 3.17cm (bande X), une distance $D_{d,1}$ de 4000m et un bâtiment de 40m, la phase de la somme des contributions des échos doubles correspond à une hauteur de -0.5cm environ. Dans ces conditions d'acquisition, l'écho diédrique est théoriquement à l'altitude du sol.

2.4.4.2.2. Modèle dans le cas pseudo-diédrique

Le bâtiment n'est plus parallèle à l'azimut mais les rugosités du sol ou sur le mur du bâtiment permettent d'obtenir un double rebond de l'onde RADAR vers le capteur [Simonetto 2002].

Considérons le trajet où la réflexion miroir a lieu sur le sol (point vert sur la figure 32a). Si le deuxième rebond de l'onde RADAR est une rétrodiffusion, alors l'onde est rétrodiffusée dans toutes les directions et ce cas revient au modèle diédrique. La contribution de ce trajet est néanmoins plus faible car moins d'énergie est renvoyée vers le capteur. Si le deuxième rebond est une réflexion spéculaire (cf. figure 32b), le capteur ne reçoit pas l'onde réfléchie et il n'y a pas d'écho sur l'image correspondant à ce trajet.

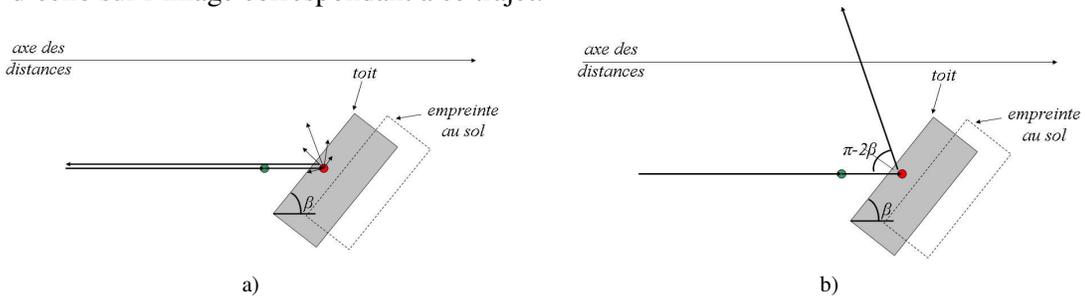


Figure 32 – Trajet effectué par l'onde RADAR lors d'un double rebond avec réflexion miroir sur le sol et a) une rétrodiffusion, b) une réflexion spéculaire dans le cas d'un pseudo-dièdre

Le trajet, lorsque la réflexion miroir a lieu sur le mur du bâtiment, est maintenant considéré. Si le mur du bâtiment est rugueux, cela revient à considérer le cas diédrique. Si le sol est lisse, l'onde n'est pas renvoyée au capteur. Le cas où le sol est rugueux et le mur est lisse est représenté en vue de coté et du dessus sur la figure 33 en considérant l'onde plane.

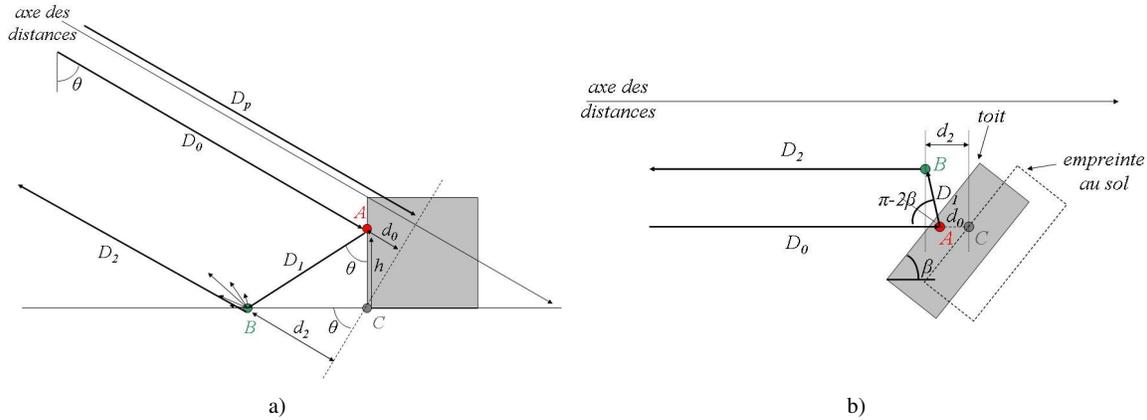


Figure 33 – Trajet effectuée par l'onde RADAR lors d'un double rebond avec réflexion miroir sur le mur
a) vue de côté, b) vue du dessus

La distance du capteur au bas de la façade est notée D_p . Les distances capteur / mur à la hauteur h et capteur / sol sont notées respectivement D_0 et D_2 . En utilisant l'hypothèse d'onde plane, D_p s'exprime par :

$$D_p = D_0(h) + d_0(h) = D_2(h) + d_2(h)$$

avec :

$$d_0(h) = h \cos(\theta) \text{ et } d_2(h) = \frac{h \sin^2(\theta)}{\cos(\theta)} \cos(\pi - 2\beta)$$

De plus :

$$D_1(h) = \frac{h}{\cos(\theta)}$$

La demi-distance parcourue par l'onde ayant effectué un double rebond à la hauteur h s'écrit :

$$D_{double}(h) = \frac{1}{2}(D_0(h) + D_1(h) + D_2(h))$$

$$D_{double} = D_p + \frac{1}{2}(D_1(h) - d_0(h) - d_2(h))$$

$$D_{double} = D_p + h \frac{1 + \sin^2(\theta) \cos(2\beta) - \cos^2(\theta)}{2 \cos \theta}$$

$$\Delta_p(h) = D_{double} - D_p = h \frac{\sin^2(\theta) \cos^2(\beta)}{\cos \theta} \quad [\text{Simonetto 02}]$$

La différence de parcours entre les rebonds doubles est maximale lorsque l'onde atteint le haut du bâtiment (h_B). Les échos doubles ne sont plus tous localisés mais s'étendent entre

$$D_p \text{ et } D_p + h_B \frac{\sin^2(\theta) \cos^2(\beta)}{\cos \theta}.$$

Pour un angle d'incidence de 45° et des hauteurs h_B de 5m, 10m, 20m et 40m, la différence de parcours entre le trajet simple et le double rebond est tracée en fonction de l'orientation du bâtiment sur la figure 34. Au premier ordre, cette différence tend vers zéro lorsque β tend vers 90° . Lorsque le bâtiment est parallèle à l'azimut ($\beta=90^\circ$), $\Delta_p(h)$ est nul au premier ordre et un développement limité à l'ordre 2 est nécessaire pour retrouver le résultat du cas diédrique (p47).

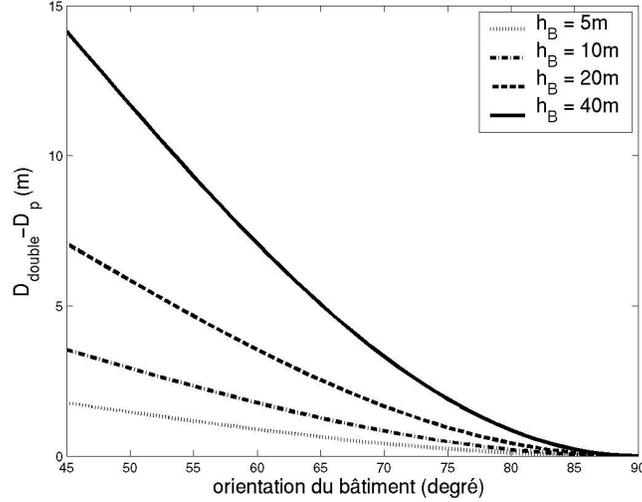


Figure 34 – Localisation des échos doubles dans le cas d'un pseudo-dièdre

Les contributions apparaissant dans la même cellule de résolution que le trajet simple vérifient :

$$\left| h \frac{\sin^2(\theta) \cos^2(\beta)}{\cos \theta} \right| \leq \frac{res_{dist}}{2} \quad (59)$$

Soit h_0 la hauteur maximale vérifiant la relation (59). h_0 s'exprime par :

$$h_0 = \frac{res_{dist}}{2} \frac{\cos \theta}{\sin^2(\theta) \cos^2(\beta)}$$

Sur la figure 35, h_0 est représentée en fonction de l'orientation du bâtiment.

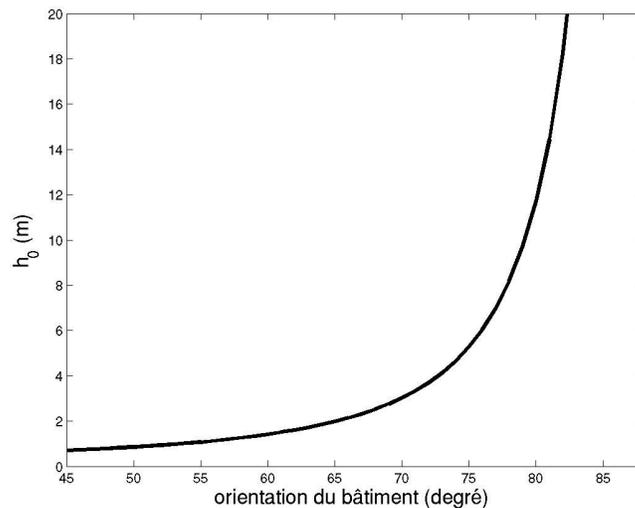


Figure 35 – hauteur maximale de la réflexion sur le mur des contributions apparaissant dans la même cellule de résolution que le trajet simple

A 50 cm de résolution et pour un bâtiment à 45°, les rebonds ayant lieu jusqu'à 0.8m de la façade apparaissent dans la même cellule de résolution que le pseudo-dièdre. Plus le bâtiment est parallèle à l'axe azimut, plus h_0 est grand. La divergence de cette courbe, lorsque le bâtiment est parallèle à l'axe azimut, ($\beta=90^\circ$) tente à montrer que toutes les contributions apparaissent dans la même cellule de résolution. Ceci est faux en très haute résolution car, pour des β proche de 90° , il faut tenir compte des termes du deuxième ordre.

Le vecteur interférométrique, résultant de l'ensemble des contributions apparaissant dans la cellule de résolution du pseudo-dièdre, s'écrit :

$$\exp\left(-j\frac{2\pi}{\lambda}(D_{p,1}-D_{p,2})\right)\left[\int_0^{h_{0,1}} A_1(h)\exp\left(-j\left(\frac{2\pi}{\lambda}\Delta_{p,1}(h)+\varphi_1(h)\right)\right)dh\right] \left[\int_0^{h_{0,2}} A_2(h)\exp\left(j\left(\frac{2\pi}{\lambda}\Delta_{p,2}(h)+\varphi_2(h)\right)\right)dh\right]$$

où φ_1 et φ_2 sont les phases, considérées aléatoires, dues au trajet de l'onde dans les cellules de résolution lors des rétrodiffusions. Il est supposé également que $\varphi_1(h) = \varphi_2(h)$.

L'argument du produit des deux intégrales fournit la hauteur des contributions ayant effectuée un double rebond et localisée dans la cellule de résolution du pseudo-dièdre. De la même façon que pour l'écho diédrique, les deux intégrales sont discrétisées de façon à respecter le théorème de Shannon. Les amplitudes des contributions $A_1(h)$ et $A_2(h)$ sont supposées identiques le long de la façade et notées A. Le produit des deux intégrales s'exprime par :

$$A^2 \left[\sum_{k_1=0}^{n_{B,1}} \exp\left(-j\left(\frac{2\pi}{\lambda}\Delta_{p,1}(k_1 \text{ pas}) + \varphi_1(k_1 \text{ pas})\right)\right) \right] \left[\sum_{k_2=0}^{n_{B,2}} \exp\left(j\left(\frac{2\pi}{\lambda}\Delta_{p,2}(k_2 \text{ pas}) + \varphi_2(k_2 \text{ pas})\right)\right) \right] \quad (60)$$

où $h_{0,1} = n_{B,1} \cdot \text{pas}$ et $h_{0,2} = n_{B,2} \cdot \text{pas}$.

Soit n_B et m_B tel que :

$$n_B = \min(n_{B,1}, n_{B,2}) \text{ et } m_B = \max(n_{B,1}, n_{B,2})$$

L'expression (60) peut être décomposée en trois termes :

$$A^2 \left[\sum_{k_1=0}^{n_B} \exp\left(-j\frac{2\pi}{\lambda}(\Delta_{p,1}(k_1 \text{ pas}) - \Delta_{p,2}(k_1 \text{ pas}))\right) \right] + \left[\sum_{\substack{k_1=0 \\ k_2 \neq k_1}}^{n_B} \sum_{k_2=0}^{m_B} \exp\left(-j\left(\frac{2\pi}{\lambda}(\Delta_{p,1}(k_1 \text{ pas}) - \Delta_{p,2}(k_2 \text{ pas})) + \varphi_1(k_1 \text{ pas}) - \varphi_2(k_2 \text{ pas})\right)\right) \right] + \left[\sum_{k_1=n_B+1}^{m_B} \sum_{k_2=0}^{m_B} \exp\left(-j\left(\frac{2\pi}{\lambda}(\Delta_{p,1}(k_1 \text{ pas}) - \Delta_{p,2}(k_2 \text{ pas})) + \varphi_1(k_1 \text{ pas}) - \varphi_2(k_2 \text{ pas})\right)\right) \right]$$

Les deux derniers termes sont des doubles sommes d'exponentielles complexes d'une variable aléatoire. En moyenne, ces doubles sommes sont nulles.

Pour une acquisition identique à celle du cas diédrique (porteur à 3000m d'altitude, $B=B_z=50\text{cm}$, $\lambda=3.17\text{cm}$ et une distance $D_{p,l}$ de 4000m) la hauteur de l'écho double par rapport au pseudo-dièdre est représenté sur la figure 36.

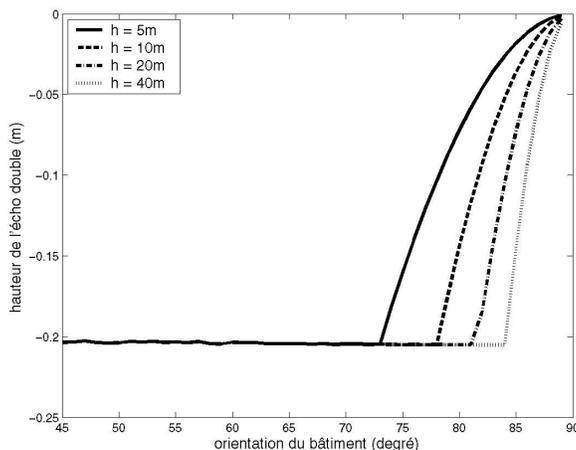


Figure 36 – hauteur de l'écho double dans le cas pseudo-diédrique pour une résolution de 50cm

Lorsque l'orientation est proche de 45° , toutes les contributions sur la façade n'apparaissent pas dans la cellule de résolution de l'écho double. La hauteur obtenue est non nulle. La distance parcourue par les contributions ayant effectué un double rebond est plus grande que la distance capteur / pseudo-dièdre. Par conséquent la hauteur résultante est négative. De plus elle est constante car le bâtiment est suffisamment grand (de hauteur supérieure à h_0) pour que les contributions apparaissent sur toute la moitié de la cellule de résolution représentée en rouge sur la figure 37a. La hauteur de l'écho double est alors reliée à la taille de la cellule de résolution

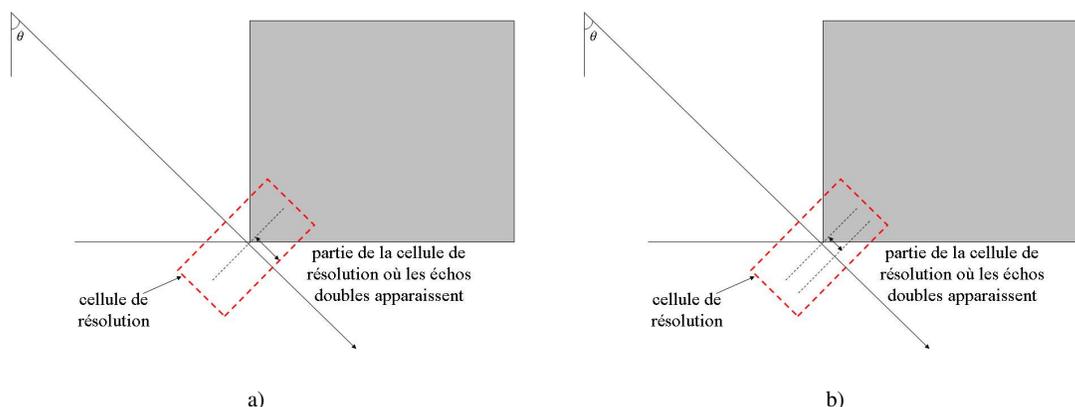


Figure 37 – localisation des échos doubles à l'intérieur de la cellule de résolution centrée sur le pseudo-dièdre lorsque la hauteur du bâtiment est a) supérieure et b) inférieure à h_0

Lorsque l'orientation du bâtiment augmente, une rupture de hauteur apparaît sur les courbes parce que le bâtiment est trop petit pour que les échos doubles apparaissent sur toute la moitié de la cellule de résolution (cf. figure 37b). La hauteur résultante est plus faible. Le cas diédrique, correspondant à une orientation de 90° , est un cas limite où il faudrait que le bâtiment ait une hauteur infini (en ne considérant que les termes du premier ordre pour le calcul de Δ_p) pour que les échos doubles apparaissent sur toute la demi-cellule de résolution.

A 50cm de résolution, les échos doubles apparaissant dans la cellule de résolution du pseudo-dièdre, sont au maximum à 20 cm au-dessous du sol et peuvent être considérés au sol.

2.4.4.2.3. Exemples

Sur la figure 38, des exemples de bâtiments avec des échos doubles sont présentés. Ces bâtiments ne sont pas parallèles à l'axe des azimuts (vertical).

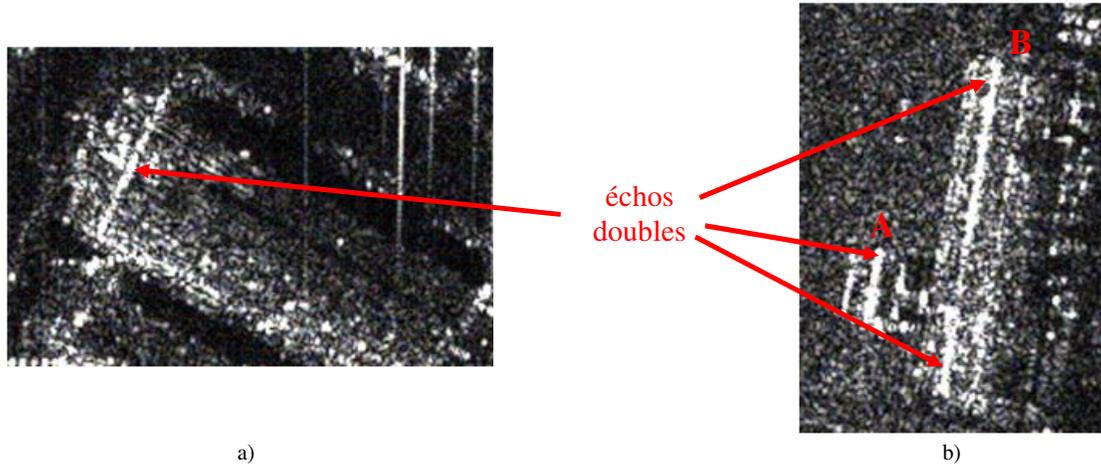


Figure 38 – Exemples d'échos doubles

Sur la figure 38b, l'écho double du grand bâtiment (B) est coupé en deux parties. En effet, à cause du petit bâtiment situé devant (A), les réflexions doubles sur le sol ne peuvent pas avoir lieu. L'accumulation des échos dans la case distance est donc plus faible voir nulle ce qui engendre la disparition de l'écho double. Les profils en amplitude selon l'axe des distances des échos de la figure 38b sont montrés sur la figure 39.

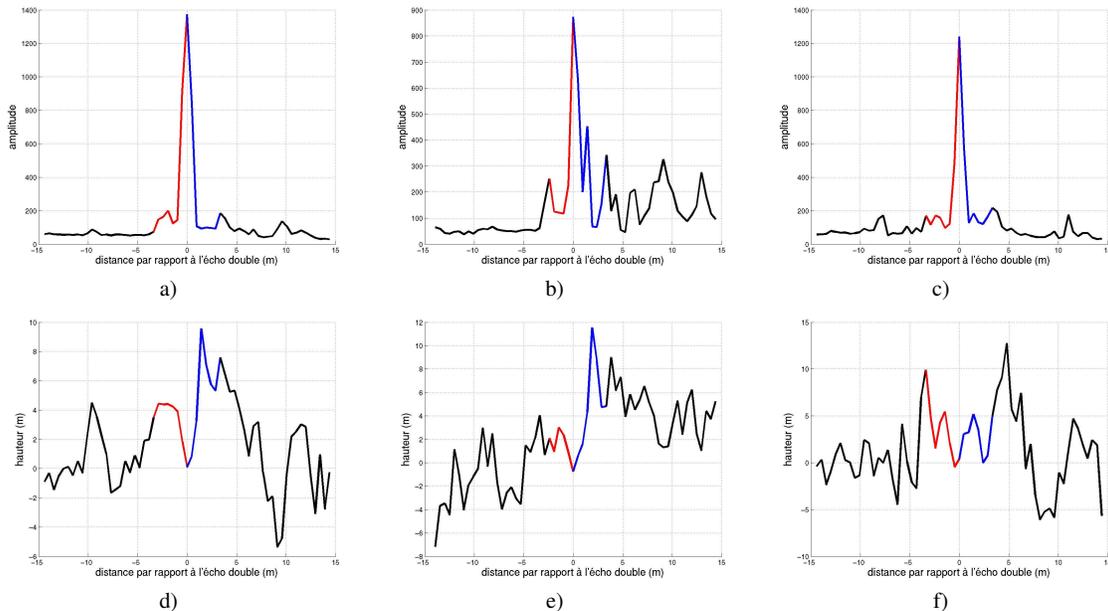


Figure 39 – Profils d'amplitude (a, b, c) et de hauteur (d, e, f) selon l'axe des distances des segments correspondant aux échos doubles B (partie supérieure), A et B partie inférieure.

En rouge et bleu sont représentées les distances correspondant respectivement aux façades et aux toits.

En rouge apparaît la zone de repliement et en bleu sont représentées les distances correspondant aux toits des bâtiments. Les échos doubles apparaissent avec une amplitude plus grande que les textures ou les échos simples. Les altitudes de ces trois échos doubles sont respectivement 0.09m, -0.74m et 0.37m avec des écart-types respectivement de 2.97m, 2.39m, 3.01m. Sur ces exemples, l'écho double est au niveau du sol.

2.4.4.2.4. Validation

Afin de valider le modèle, nous nous proposons de déterminer la hauteur par rapport au sol de 18 échos doubles. Chaque écho est sélectionné manuellement et l'altitude interférométrique est calculée sur l'interférogramme multivue 2x2 car les images sont sur-échantillonnées d'un facteur 2 selon les deux axes. Soit n le nombre de cellule de résolution où apparaît l'écho. La phase interférométrique et la cohérence associées sont notées φ_i et γ_i . La phase interférométrique de l'écho est calculée par :

$$\varphi_{\text{écho}} = \arg \left(\sum_{i=1}^n \gamma_i \exp(j\varphi_i) \right)$$

De cette façon, les cellules de résolution fortement cohérentes qui fournissent une information altimétrique précise, sont privilégiées dans le calcul de la phase interférométrique globale. L'altitude du sol est déterminée sur une zone dégagée proche de l'écho double par croissance de région sur l'interférogramme multivue 8x8. La vérité terrain sur la hauteur des bâtiments est donnée par la BDTopo^{©IGN} lorsqu'elle est disponible ou par mesure monoscopique (cf §2.3.1).

Sur la figure 40 est représentée la différence d'altitude entre le sol et l'écho double en fonction de la hauteur monoscopique de la façade.

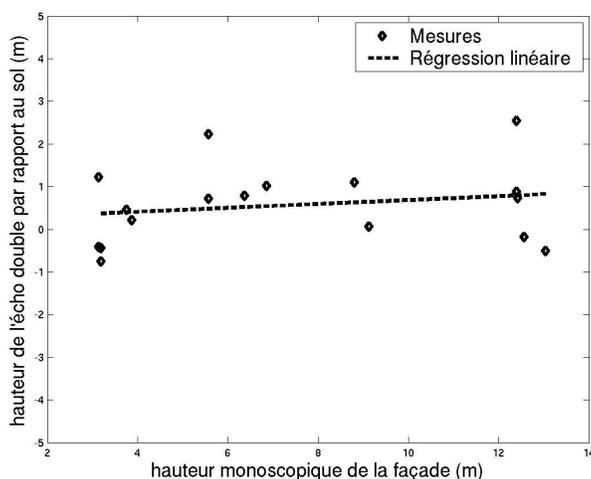


Figure 40 – hauteur de l'écho double par rapport au sol

L'interpolation par les moindres carrés des échantillons donne une droite de pente quasi-nulle (pente de 4.5cm / m). L'influence de la hauteur de la façade sur l'altitude de l'écho double est donc négligeable.

Néanmoins, il y a un biais de 50cm environ entre les altitudes du sol et des échos doubles. Les écart-types interférométriques sur les mesures du sol et des échos varient entre 0.75m et 3m. Le biais est inférieur aux écarts-types sur les mesures et nous considérerons dans la suite de ce document que l'écho double se situe à la même altitude que le sol.

2.4.4.2.5. Conclusion

L'écho double est constitué d'une somme d'échos apparaissant dans la même cellule de résolution. Plus le bâtiment est élevé et parallèle à la trajectoire du porteur et plus le nombre de contribution dans la cellule de résolution du pseudo-dièdre est important. Il a donc une

amplitude généralement supérieure aux échos simples. L'utilisation du modèle proposé par [Simonetto 2002] en interférométrie a montré que l'altitude de l'écho double est celle du sol dans les conditions d'acquisition similaires à celles des images présentées en annexe A.

2.4.4.3. Les échos triples

2.4.4.3.1. Modèles

Les échos triples subissent trois réflexions/rétrodiffusions. De même que pour l'écho double, ils peuvent être différenciés suivant le type de surface où a lieu la première réflexion ou rétrodiffusion (sol ou mur). Pour les deux cas, l'angle formé entre l'axe des distances et le bâtiment est noté β . La distance entre le coin, formé par la façade et le sol, et le capteur est notée D_p . D_0 représente la distance du capteur à l'endroit où a lieu le premier rebond.

Sur la figure 41 sont représentés le trajet mur / sol / mur pour une vue en coupe et du dessus du bâtiment.

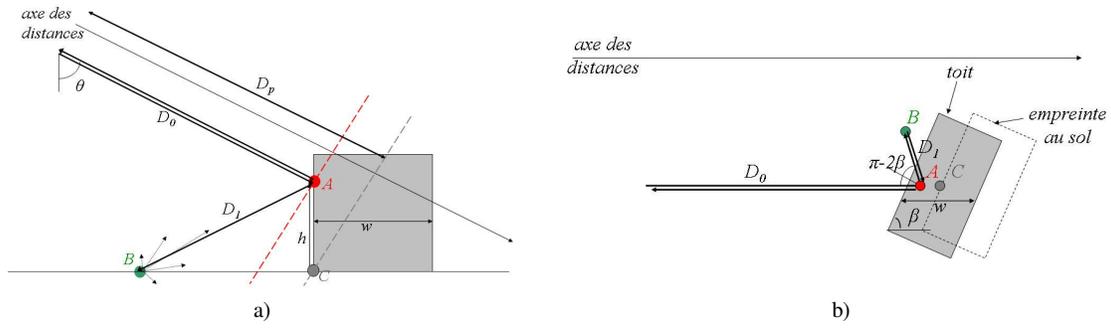


Figure 41 – Trajet mur/sol/mur effectués par les échos triples a) vue en coupe, b) vue du dessus

Dans le cas des échos triples mur / sol / mur, D_0 représente la distance entre le capteur et le mur et D_1 correspond à celle entre le mur et le sol. Pour une réflexion mur / sol / mur à la hauteur h du bâtiment, en utilisant l'hypothèse d'onde plane, la demi-distance parcourue en tout par l'onde est notée D_{msm} et s'exprime par :

$$D_{msm}(h) = D_0(h) + D_1(h)$$

En prenant D_p comme distance de référence, D_0 et D_1 s'expriment par :

$$D_0(h) = D_p - h \cos(\theta) \text{ et } D_1(h) = \frac{h}{\cos(\theta)}$$

D'où

$$D_{msm}(h) = D_p + \frac{h}{\cos(\theta)} - h \cos(\theta) \quad (61)$$

La distance où sont situés les échos triples mur / sol / mur ne dépend pas de l'orientation du bâtiment.

Tout se passe comme si c'était le symétrique du point B par rapport à la façade qui répondait. Ce symétrique se situe sur le sol. Par conséquent la hauteur interférométrique d'un écho triple mur/sol/mur est l'altitude du sol indépendamment de la hauteur sur la façade à laquelle la réflexion a eu lieu.

Pour les réflexions sol / mur / sol le trajet est représenté selon les deux même vues que précédemment sur la figure 42. L'onde subit une réflexion miroir sur le sol puis sur la façade du bâtiment pour être ensuite rétrodiffusée vers le capteur. Pour que ce trajet soit possible, il faut que le bâtiment présente un rebord (gouttière, corniche, balcon...) à la hauteur h . D_0 représente, dans ce cas, la distance entre le capteur et le sol où a lieu le premier rebond. D_1 correspond à la distance entre le rebond et le mur.

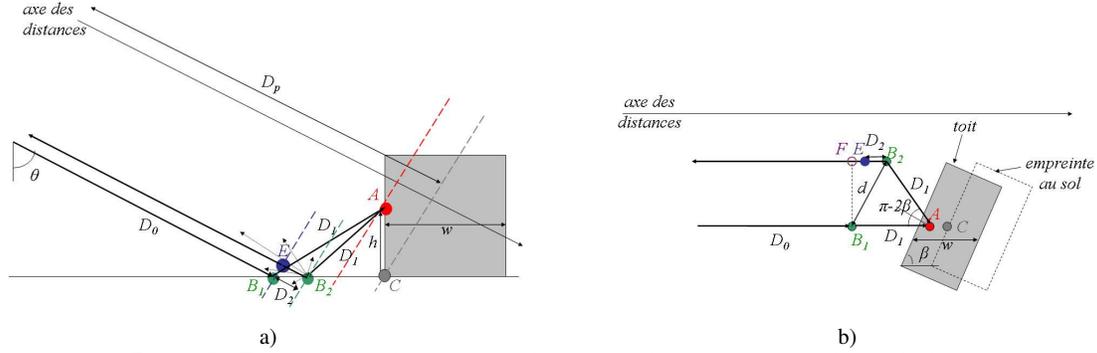


Figure 42 – Trajet sol/mur/sol effectués par les échos triples a) vue en coupe, b) vue du dessus

En supposant le front d'onde plan, les distances capteur- B_1 et capteur- E sont identiques et notées D_0 . Pour une réflexion triple à la hauteur h du bâtiment, la demi-distance parcourue par l'onde est notée D_{sms} :

$$D_{sms} = D_0 + D_1 + \frac{1}{2} D_2$$

La distance D_1 vaut :

$$D_1 = \frac{h}{\cos(\theta)}$$

La distance d entre les points B_1 et B_2 appartenant au sol s'exprime par :

$$d = 2D_1 \cos(\beta)$$

La distance d' représente la distance entre les points B_2 et le point F , projeté en géométrie terrain du point E .

$$d' = d \cos(\beta)$$

D'où

$$D_2 = d' \sin(\theta)$$

$$D_2 = 2h \tan(\theta) \cos^2(\beta)$$

De plus, en considérant le symétrique du point A par rapport au sol on a :

$$D_0 + D_1 = D_p + h \cos(\theta)$$

D'où

$$D_{sms} = D_p + h \cos(\theta) + h \tan(\theta) \cos^2(\beta) \quad (62)$$

Lorsque le mur est rugueux par rapport à la longueur d'onde, l'onde ne subit pas une réflexion miroir mais une rétrodiffusion et β est considéré comme égal à $\pi/2$.

Lors d'un trajet sol / mur / sol, l'onde atteint un point virtuel situé à une hauteur négative. La hauteur de l'écho triple est donc l'opposé de la hauteur de la façade où a eu lieu la réflexion.

Pour les deux types d'échos triples, le chemin parcouru par l'onde électromagnétique est donc plus long que celui des réflexions simples et doubles. Ils apparaissent alors plus éloignés du capteur que les échos doubles. Les échos triples ne sont pas tous localisés dans une même cellule de résolution. Ils ont généralement une amplitude plus faible que les échos simples car le signal électromagnétique a perdu de l'énergie lors des multiples réflexions/rétrodiffusions.

Soit w la largeur du bâtiment selon l'axe des distances et h_B la hauteur du bâtiment. D_{bord_toit} représente la distance entre le capteur et le bord du toit le plus éloigné du capteur (cf. figure 43). Les distances maximales à laquelle apparaissent les deux types d'échos sont représentées. Il est supposé, dans la suite de cette thèse, qu'ils ne sont visibles sur les images en amplitude seulement si aucun autre signal n'est reçu par le capteur, c'est-à-dire s'ils apparaissent dans les ombres ou si le toit est non-rétrodiffusant.

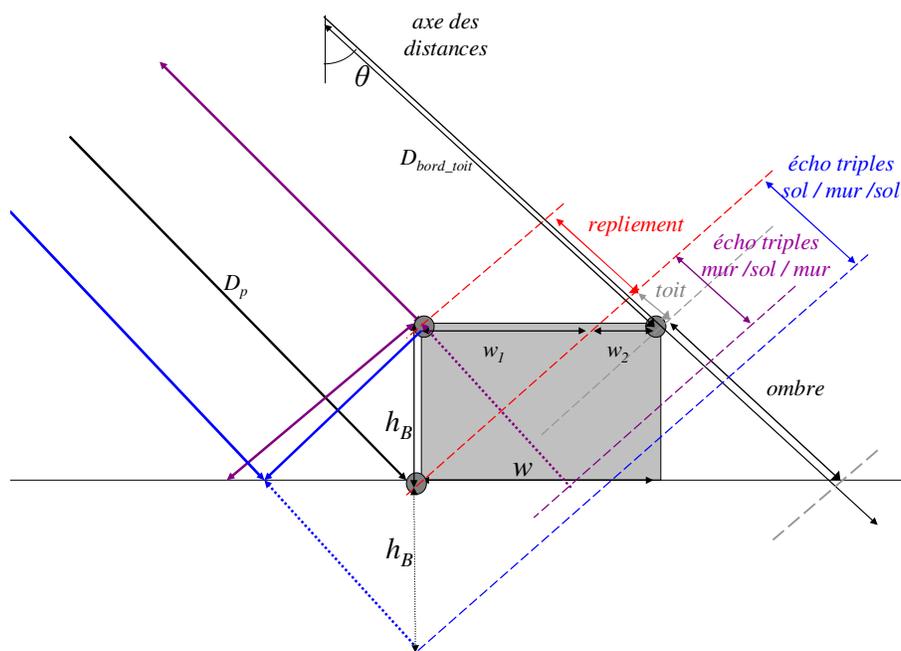


Figure 43 – localisation des échos triples

Soient w_1 et w_2 tel que :

$$w = w_1 + w_2$$

En utilisant l'hypothèse d'onde plane, D_{bord_toit} s'exprime en fonction de D_p par :

$$D_{bord_toit} = D_p + (w - w_1) \sin(\theta)$$

$$D_{bord_toit} = D_p + w \sin(\theta) - h_B \cos(\theta)$$

L'écho triple mur / sol / mur le plus éloigné atteint la façade à la hauteur h_B . Il apparaît dans l'ombre si :

$$D_{msm}(h_B) > D_{bord_toit}$$

$$h_B > w \cos(\theta) \sin(\theta)$$

L'écho triples sol / mur / sol le plus éloigné atteint également la façade à la hauteur h_B . Il apparaît dans l'ombre si :

$$D_{sms}(h_B) > D_{bord_toit}$$

$$h_B > \frac{w \cos(\theta) \sin(\theta)}{2 \cos^2(\theta) + \sin(\theta) \cos^2(\beta)} \quad (63)$$

Le rapport w/h_B est tracé sur la figure 44 pour déterminer la limite visibilité des échos triples dans les ombres, selon l'angle d'incidence, le type de réflexions et l'orientation du bâtiment. La zone en dessous de chaque courbe correspond aux rapports w/h_B pour lesquels les échos sont visibles dans l'ombre. Pour un angle d'incidence inférieur à 45° , les échos triples sol / mur / sol sont théoriquement plus éloignés que les échos mur / sol / mur ; et ceci pour toutes les orientations des bâtiments. Pour une hauteur de réflexion sur le mur fixe, plus l'incidence est rasante et plus l'orientation du bâtiment influe sur la distance parcourue par l'onde car le trajet mur / sol (seul terme dépendant de β) est important.

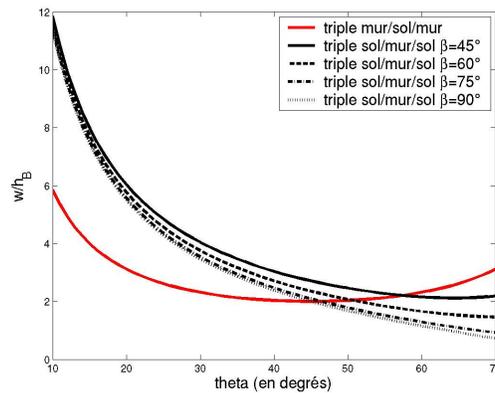


Figure 44 – Visibilité des échos triples dans les ombres en fonction du rapport w/h_B du bâtiment

2.4.4.3.2. Validation

La figure 45 présente deux exemples d'images RSO (multivue 2x2) de bâtiments, d'orientations différentes (90° et 60°), où des échos triples sont visibles. L'angle d'incidence de l'onde RADAR étant inférieur à 45° , seul les échos triples sol/mur/sol sont validés. Les signatures des bâtiments ont été choisies de telle sorte que l'écho double et l'écho triple soient visibles.

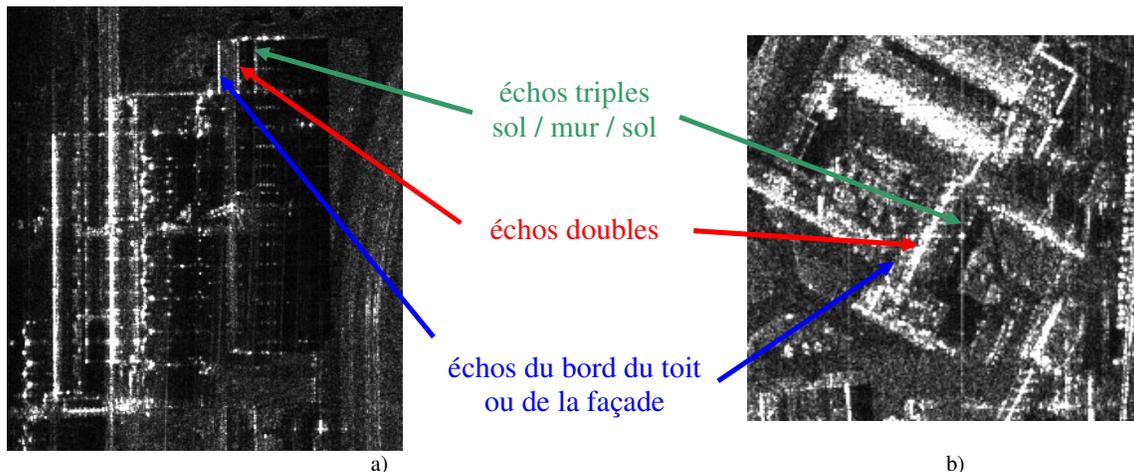


Figure 45 – Exemples de signatures de bâtiments avec des échos triples pour une orientation de a) 90° et b) 60°

Sur la figure 46 sont représentés les profils en amplitude et en hauteur de l'écho double selon l'axe des distances pour l'image de la figure 45a. L'abscisse 0 de ce profil correspond à la distance où apparaît l'écho double. Cet écho sert de référence pour le calcul de hauteur. Sur le profil interférométrique sont représentés en rouge les écarts-types altimétriques des points de l'interférogramme utilisés (bord du toit, écho double, écho triple).

Dans ce premier exemple, $\beta = 90^\circ$. Sur les profils de la figure 46, la distance entre l'écho double et l'écho triple est de 6.7m. La BDTopo^{©IGN} n'est pas disponible sur ce bâtiment. Par mesure monoscopique, la hauteur de l'écho simple peut être estimée à moins d'un mètre de précision (cf. §2.3.1) en tenant compte de la distance entre l'écho double et l'écho simple. Dans le cas présent, la hauteur du bâtiment est estimée à 8.1m. En utilisant (62), la distance théorique des échos triples sol / mur / sol avec réflexion sur le mur à cette hauteur est de 6.7m. La hauteur des échos triples est lue sur l'interférogramme et vaut -8m. Le modèle est donc en accord avec les données.

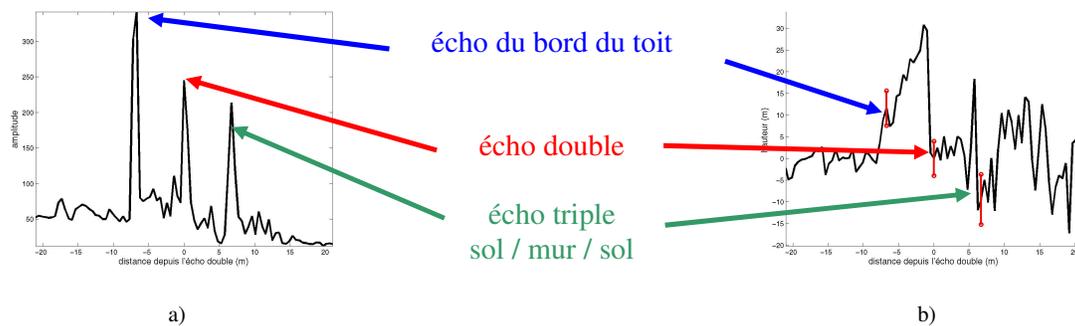


Figure 46 – Profils a) amplitude et b) interférométrique centrés sur l'écho double

La figure 45b présente un cas où la réflexion sur le mur est une réflexion miroir. Le bâtiment est orienté selon un angle $\beta = 60^\circ$. La hauteur du bâtiment, fournie par la BDTopo^{©IGN} est de 9m. Les profils amplitude et hauteur interférométrique sont présentés en figure 47. La mesure interférométrique de l'écho de la façade donne une hauteur de 6.5m. La mesure monoscopique sur le pic d'amplitude à -5.75m permet d'estimer la hauteur de la façade à 7.65m, ce qui laisserait à penser que cet écho ne correspond pas au haut du toit mais plutôt à un élément de la façade légèrement en dessous. En considérant la relation (62), un écho triple ayant atteint le mur à une hauteur de 9m devrait apparaître à une distance de 8.73m. Sur le profil en amplitude, un pic d'énergie à une distance de 9.3m est visible. Sachant que l'incertitude sur les données de la BDTopo^{©IGN} est de 1m, la distance théorique des échos triples doit être comprise entre 7.76m et 9.70m. Cet écho triple présente une hauteur, lue sur l'interférogramme, de -8.3m mais l'écart-type sur cette mesure est très important car l'écho a une amplitude moyenne plus faible que dans le cas précédent. Sur cette image, le modèle correspond encore aux données.

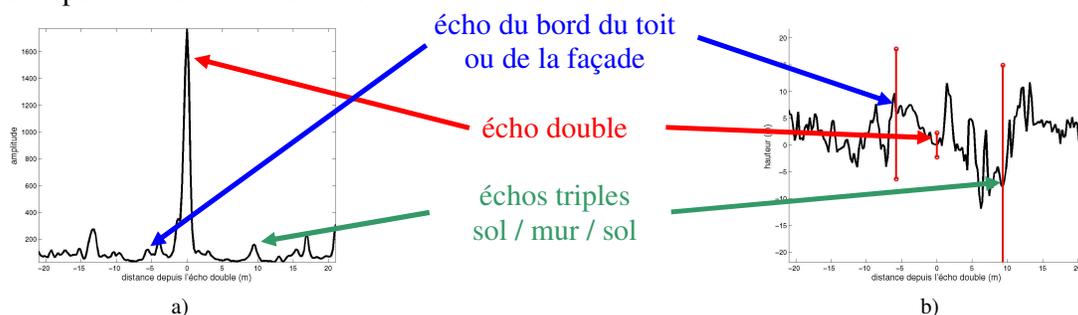


Figure 47 – Profils a) amplitude et b) interférométrique centrés sur l'écho double

La figure 48a présente une signature où, malgré l'orientation du bâtiment, l'onde a subi une rétrodiffusion et non une réflexion. En effet la distance entre l'écho simple et l'écho double est de 10m (cf. figure 48c) et, par mesure monoscopique, la hauteur calculée est de 13m. Lorsque β est considéré comme valant 90° la distance théorique entre l'écho double et l'écho triple est de 10m, ce qui est vérifié sur le profil d'amplitude et sur l'interférogramme où la hauteur lue est de -14.9m (figure 48d).

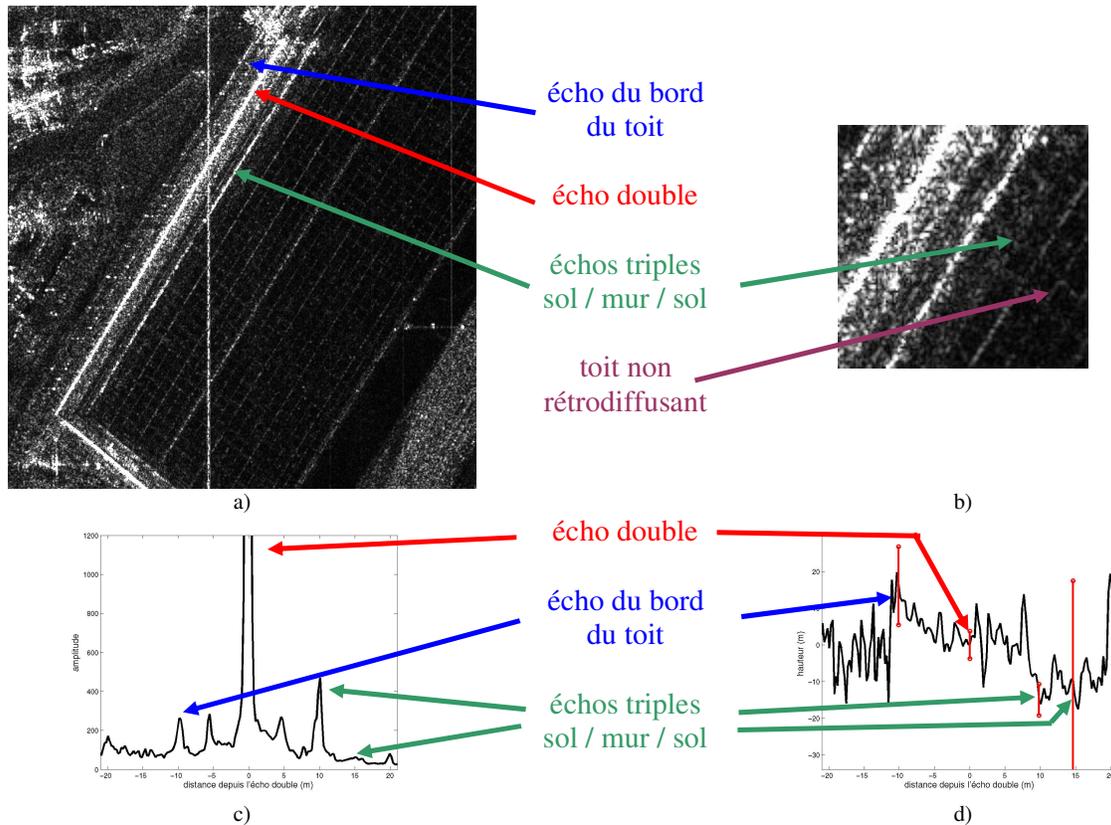


Figure 48 – a) Exemples d'échos triples sol/mur/sol sans réflexion miroir sur la façade du bâtiment profil b) amplitude et c) interférométrie de l'écho double

De plus lors de la rétrodiffusion (inhabituellement forte) sur le mur, une partie de l'énergie est émise dans les autres directions, notamment celle correspondant à la réflexion miroir (50°). β pourrait être considéré non plus comme un angle mais comme un intervalle centré autour de 90° . Le trajet, décrit dans le cas des coins réflecteurs, serait donc emprunté par une faible partie de l'énergie, ce qui expliquerait le phénomène est visible sur la figure 48b : après l'écho triple linéaire, il y a une zone où des points peu brillants apparaissent (cf. figure 48b). Ces points peu brillants sont visibles car le toit est peu rétrodiffusant. Ces échos triples se traduisent par une amplitude localement plus importante à 15m de l'écho double. Théoriquement, la position des échos triples ayant subi une réflexion miroir sur la façade est de 14.22m. La hauteur lue à cette position est de -7m mais l'écart-type sur cette mesure est très important. Des échos triples apparaissent plus loin, jusqu'à 16m, ce qui correspondrait à des angles miroirs « fictifs » inférieurs à 50° .

2.4.4.3. Conclusion

La présence sur les images RAMSES en bande X d'échos triples sol / mur / sol apparaissant dans les ombres a été vérifiée. De part l'angle d'incidence, les échos triples mur / sol / mur ne sont pas visibles et n'ont pas pu être vérifiés. Les échos triples sol / mur / sol ont un niveau de rétrodiffusion faible et une hauteur négative. Néanmoins selon le type de réflexion sur la façade, il est difficile de prévoir la localisation de ces échos. En revanche, l'information altimétrique de ces échos permet d'avoir une information sur la hauteur minimale de la façade où a eu lieu la réflexion.

2.4.4.4. Et plus encore...

En milieu urbain, la complexité de certains bâtiments peuvent faire apparaître des échos ayant subi plus de trois réflexions / rétrodiffusions. Sur la figure 49 est présenté la signature RADAR d'un bâtiment rectangulaire avec un patio. Sur le profil d'amplitude associé à l'écho double de la façade, un pic d'amplitude est visible à -2.4m, ce qui correspond à une mesure monoscopique de 3.2m. La mesure monoscopique sur les ombres ainsi que la BDTopo^{©IGN} fournissent une estimation de la hauteur à 5m. Un pic d'énergie correspondant à la fin du toit présente une hauteur interférométrique de 5.6m. L'orientation du bâtiment est de 80°. Par conséquent les échos triples devraient se situés à 3.9m de l'écho double mais ils sont cachés par la rétrodiffusion du toit. Néanmoins, un écho apparaît dans les ombres à une distance de 9.4m. La hauteur interférométrique lue est de -7m. Cet écho est sûrement issu de multiples réflexions à l'intérieur du patio.

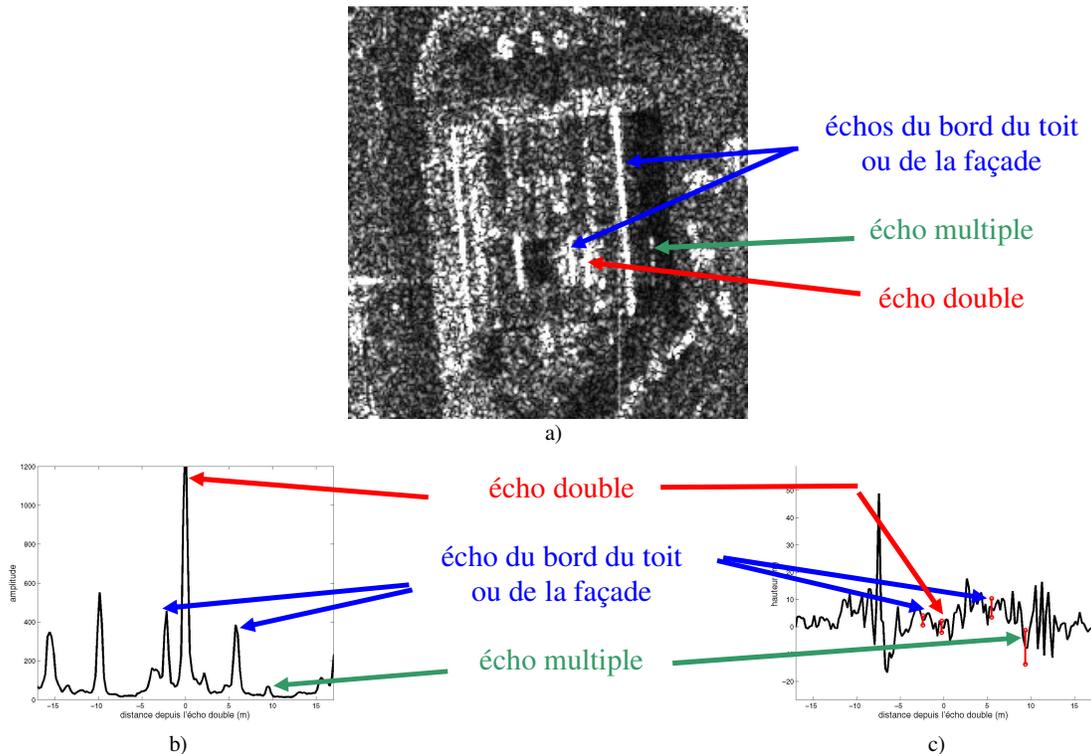


Figure 49 – a) Exemples d'échos triples sol/mur/sol sans réflexion miroir sur la façade du bâtiment
profils b) amplitude et c) interférométrique de l'écho double

2.4.5. Conclusion

La phénoménologie des images RSO haute résolution en bande X. et en milieu urbain a été présentée. Les textures des images RSO sont liées à la longueur d'onde de la fréquence centrale utilisée et aux irrégularités des surfaces observées. Des échos forts apparaissent sur les images lorsque l'onde RADAR est renvoyée de façon directive vers le capteur. L'utilisation d'un signal énergétique engendre des phénomènes de réflexions multiples. Les échos diédriques apparaissent, en imagerie RSO haute résolution, dans la même case distance que le dièdre formé par le sol et la façade du bâtiment. Lorsque le bâtiment n'est pas parallèle à la trajectoire du porteur, les échos doubles sont étalés spatialement et n'apparaissent pas systématiquement dans la cellule de résolution du pseudo-dièdre. A 50cm de résolution, dans le cas dièdre comme pseudo-dièdre, ils sont à l'altitude du sol. Les échos triples apparaissent dans les ombres et peuvent fournir une information altimétrique supplémentaire sur le bâtiment.

2.5. Etude des mélanges d'informations 3D au sein d'une zone de repliement

2.5.1. Origine du repliement

Le repliement est un phénomène engendré par l'acquisition en distance et en visée latérale, du système imageur et apparaît lorsque l'angle entre une partie de objet et le sol est supérieur à l'angle d'incidence de l'onde RADAR. Plusieurs zones se situant à la même distance du capteur rétrodiffusent l'onde RADAR simultanément. L'onde reçue par le capteur est donc un mélange d'ondes cohérentes constitué de contributions ayant des amplitudes et des altitudes différentes. Au sein de chaque cellule de résolution d'une zone de repliement l'information altimétrique fournie par l'interférogramme n'est alors pas directement utilisable. Dans ce paragraphe, un modèle pour séparer ce mélange, à partir de données interférométriques, lorsqu'il n'y a que deux contributions (notées P_A et P_B), est proposé (cf. figure 50). Ce modèle est ensuite évalué sur l'estimation de hauteur de façades de bâtiments.

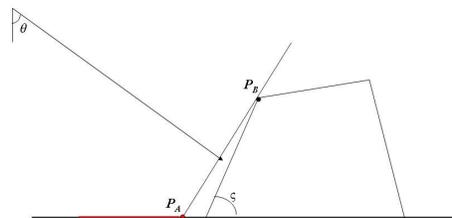


Figure 50 – Le repliement en imagerie RSO : objet avec $\zeta > \theta$

2.5.2. Modélisation d'un mélange de deux contributions d'altitudes différentes

2.5.2.1. Interférogramme dans une cellule de résolution sans repliement

La mesure interférométrique correspond à une mesure de différence de parcours entre les deux ondes RADAR rétrodiffusées par la même cellule de résolution (cf. §1.3). Cette différence de parcours est très faible et se mesure en fraction de longueur d'onde. Soit P_A une région où il n'y a pas de repliement.

L'onde RADAR rétrodiffusée par les m rétrodiffuseurs de la cellule de résolution s'écrit:

$$\sum_{i=1}^m A_i \exp(-j\varphi_{A,i}) = A \exp(-j\varphi_{A,0})$$

$\varphi_{A,0}$ est appelée phase propre et est directement reliée à la distance capteur / scène ainsi qu'au type, nombre et disposition des rétrodiffuseurs dans la cellule de résolution. Néanmoins, la phase étant mesurée entre 0 et 2π , l'information de distance est perdue. De plus la modélisation de la rétrodiffusion de l'onde au sein d'une même cellule résolution est très complexe. Ce terme est donc généralement considéré comme aléatoire spatialement (d'une cellule de résolution à l'autre). En revanche la phase propre est reproductible.

Sur l'autre image du couple interférométrique, l'angle d'incidence étant légèrement différent, le trajet parcouru par l'onde est également différent. Nous faisons l'hypothèse simplificatrice suivante : les amplitudes des pixels correspondants sur les deux images du couple interférométrique sont identiques.

Les ondes RADAR reçues s'écrivent alors :

$$\sum_{i=1}^m A_i \exp(-j(\varphi_{A,0}^i + \varphi_A)) = A \exp(-j(\varphi_{A,0} + \varphi_A))$$

où φ_A est la phase interférométrique. Elle est reliée à la différence de marche par l'équation (29), p27.

Dans des conditions adéquates d'acquisition, lorsque la base interférométrique est inférieure à la base critique, l'évolution de la phase entre les deux images est inférieure à 2π . L'interférogramme est formé de la façon suivante :

$$A_i \exp(-j\varphi_{A,0}^i) A \exp(j(\varphi_{A,0}^i + \varphi_A)) = A^2 \exp(\varphi_A^i)$$

La composante aléatoire (phase due au trajet) est supprimée et la phase interférométrique et donc l'altitude relative de la cellule de résolution est obtenue.

2.5.2.2. Modèle de mélange d'altitude au sein d'une zone de repliement

Soient n cellules de résolution où il y a du repliement. Lorsque deux régions P_A et P_B , d'altitudes différentes, apparaissent dans une même cellule de résolution, la mesure interférométrique du mélange P_C est donnée par :

$$C_i^2 \exp(j\varphi_C^i) = (A_i \exp(-j\varphi_{A,0}^i) + B_i \exp(-j\varphi_{B,0}^i)) (A_i \exp(j(\varphi_{A,0}^i + \varphi_A)) + B_i \exp(j(\varphi_{B,0}^i + \varphi_B)))$$

$$C_i^2 \exp(j\varphi_C^i) = A_i^2 \exp(j\varphi_A^i) + B_i^2 \exp(j\varphi_B^i) + 2A_i B_i \exp\left(j\frac{\varphi_A^i + \varphi_B^i}{2}\right) \cos\left(\varphi_{A,0}^i - \varphi_{B,0}^i - \frac{\varphi_A^i - \varphi_B^i}{2}\right) \quad (64)$$

De cette expression ressortent trois termes. Les deux premiers correspondent aux contributions interférométriques considérées séparément lorsqu'il n'y a pas de repliement. Le troisième correspond aux interférences entre les deux contributions.

Les phases propres sont aléatoires. Le terme en cosinus est donc aléatoire d'une cellule de résolution à l'autre. En moyennant l'expression précédente, sur les n pixels de la zone de repliement où les deux mêmes contributions apparaissent, le mélange s'écrit :

$$C^2 \exp(j\varphi_C) = A^2 \exp(j\varphi_A) + B^2 \exp(j\varphi_B)$$

où

$$A^2 = \frac{1}{n} \sum_i A_i^2 \text{ et } B^2 = \frac{1}{n} \sum_i B_i^2$$

$$\varphi_A = \frac{1}{n} \sum_i \gamma_i \varphi_A^i \text{ et } \varphi_B = \frac{1}{n} \sum_i \gamma_i \varphi_B^i \text{ avec } \gamma_i \text{ la cohérence associée au pixel } i$$

Ce modèle peut être vu comme la composition dans le plan complexe des vecteurs interférométriques des deux contributions (cf. figure 51).

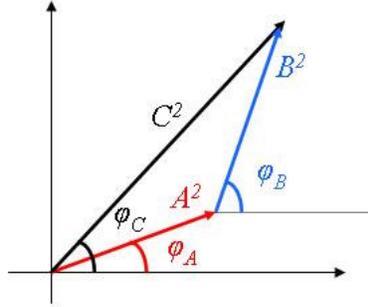


Figure 51 – Représentation du modèle de mélange dans le plan complexe

La phase interférométrique correspondant à la région P_B s'écrit :

$$\varphi_B = \text{atan} \left(\frac{C^2 \sin(\varphi_C) - A^2 \sin(\varphi_A)}{C^2 \cos(\varphi_C) - A^2 \cos(\varphi_A)} \right)$$

Le contraste, rapport des amplitudes des zones avec et sans repliement, est défini par :

$$R = \frac{C}{A}$$

D'où:

$$\varphi_B = \text{atan} \left(\frac{R^2 \sin(\varphi_C) - \sin(\varphi_A)}{R^2 \cos(\varphi_C) - \cos(\varphi_A)} \right)$$

La phase interférométrique φ_B tend vers φ_C lorsque $R \gg 1$ et φ_B tend vers φ_A lorsque $R \ll 1$. Pour de très forts ou très faibles contrastes le mélange d'altitudes provenant du repliement peut être considéré comme une altitude unique.

Il est également intéressant d'exprimer la phase φ_B en fonction de la différence $\varphi_C - \varphi_A$ qui est proportionnelle à la hauteur apparente de l'objet, notée $h_{\text{apparente}}$, dans la zone de repliement :

$$\varphi_B - \varphi_A = \text{atan} \left(\frac{R^2 \sin(\varphi_C - \varphi_A)}{R^2 \cos(\varphi_C - \varphi_A) - 1} \right)$$

La différence $\varphi_B - \varphi_A$ est proportionnelle à la hauteur de l'objet lorsque le mélange d'altitude a été séparé. Cette hauteur est notée $h_{\text{modèle}}$.

Les erreurs maximales sur la différence $\varphi_C - \varphi_A$, provenant du terme aléatoire de l'expression (64), s'écrivent :

$$\varphi_{B,\text{max}} - \varphi_A = 2 \text{atan} \left(\frac{R \sin\left(\frac{\varphi_C - \varphi_A}{2}\right)}{R \cos\left(\frac{\varphi_C - \varphi_A}{2}\right) - 1} \right) \quad \text{et} \quad \varphi_{B,\text{min}} - \varphi_A = 2 \text{atan} \left(\frac{R \sin\left(\frac{\varphi_C - \varphi_A}{2}\right)}{R \cos\left(\frac{\varphi_C - \varphi_A}{2}\right) + 1} \right)$$

2.5.3. Interprétation du modèle

2.5.3.1. Influence des paramètres

Le ratio des hauteurs interférométriques avec et sans modèle est appelé facteur correctif interférométrique et est noté f . Il peut s'exprimer en fonction de R et de la hauteur apparente de l'objet :

$$f = \frac{\varphi_B - \varphi_A}{\varphi_C - \varphi_A} = \frac{\text{atan}\left(\frac{R^2 \sin(\varphi_C - \varphi_A)}{1 + R^2 \cos(\varphi_C - \varphi_A)}\right)}{\varphi_C - \varphi_A}$$

f_{\min} et f_{\max} sont définis par :

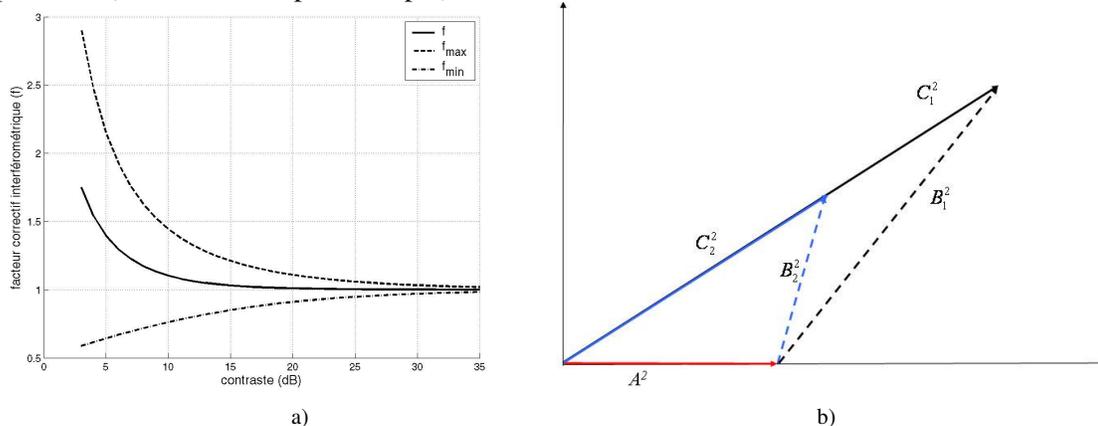
$$f_{\min} = \frac{\varphi_{B,\min} - \varphi_A}{\varphi_C - \varphi_A} \text{ et } f_{\max} = \frac{\varphi_{B,\max} - \varphi_A}{\varphi_C - \varphi_A}$$

Ce rapport est intéressant car il permet d'obtenir le rapport entre la hauteur obtenue avec modèle et la hauteur apparente sans avoir à estimer l'altitude d'ambiguïté. En effet :

$$f = \frac{\varphi_B - \varphi_A}{\varphi_C - \varphi_A} = \frac{2\pi \frac{h_{\text{modèle}}}{h_{\text{amb}}}}{2\pi \frac{h_{\text{apparente}}}{h_{\text{amb}}}} = \frac{h_{\text{modèle}}}{h_{\text{apparente}}}$$

2.5.3.1.1. Influence du contraste

R est supposé supérieur à 1. Dans le cas contraire, le modèle étant symétrique, il suffit d'interchanger les primitives A et C . A hauteur apparente fixe, f est une fonction décroissante du contraste, il est toujours supérieur à 1 (cf. figure 52a). Cela signifie que la hauteur apparente sous-estime la hauteur réelle de l'objet. En effet, la hauteur apparente provient d'un mélange entre la hauteur réelle de l'objet et la hauteur, plus faible, de la zone où a lieu le repliement (hauteur du sol par exemple).



A hauteur apparente constante, le ratio des hauteurs interférométriques est donc plus important pour de faibles contrastes où la contribution de la région P_B est noyée dans celle de P_A . En effet, sur la figure 52b, le vecteur interférométrique correspondant à la région de plus faible altitude est représenté avec une phase nulle (vecteur rouge). En noir et bleu sont représentés, pour deux contrastes, les vecteurs interférométriques des mélanges des deux contributions. Les vecteurs en traits discontinus représentent les contributions P_B pour ces

deux contrastes. La variation de la phase de P_B est visiblement plus importante lorsque le contraste est faible.

2.5.3.1.2. Influence de la hauteur apparente

Sur la figure 53, l'évolution du facteur correctif en fonction de la hauteur apparente est représentée lorsque le contraste est constant. L'influence de la hauteur apparente sur le facteur correctif interférométrique est beaucoup plus faible que l'influence du contraste. Il y a néanmoins une légère décroissance du ratio en fonction de la hauteur apparente. En effet, pour un même contraste, lorsque la hauteur apparente augmente, la contribution du sol dans le mélange diminue.

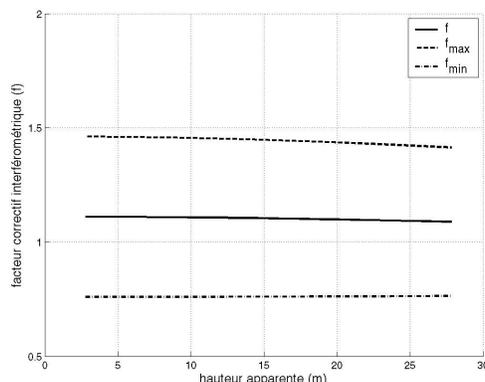


Figure 53 – Evolution du facteur correctif en fonction de la hauteur apparente $\varphi_C - \varphi_A$

2.5.3.1.3. Influence du nombre de pixel sur la zone d'estimation

Il est visible sur la figure 52 et la figure 53 que l'erreur provenant de la partie aléatoire du mélange est très importante. Pour diminuer ces erreurs il faut augmenter le nombre de cellules de résolution sur lesquelles le moyennage est effectué. Pour déterminer l'influence de ce nombre de pixels, 100 phases comprises entre 0 et 2π sont générées aléatoirement. La valeur moyenne du cosinus de ces phases est représentée sur la figure 54 en fonction de la taille en pixel du moyennage spatial. Il est visible que pour un faible nombre d'échantillons, la valeur moyenne du cosinus est très différente de 0. Dans l'exemple présenté, à partir de 50 échantillons, cette valeur moyenne est faible. Pour utiliser le modèle proposé, il faut donc un moyennage spatial le plus grand possible.

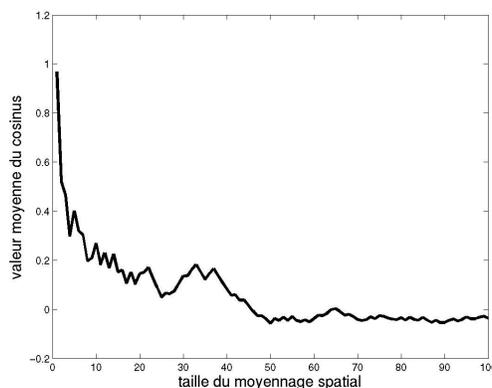


Figure 54 – Evolution de la valeur moyenne du cosinus d'une variable aléatoire en fonction du nombre de réalisations

2.5.3.2. Influence des erreurs d'estimations sur les paramètres

Le facteur correctif dépend de l'estimation de l'altitude de la zone où a lieu le repliement (φ_A), du mélange d'altitude (φ_C), et du contraste (R). Dans ce paragraphe, la robustesse de ce modèle vis-à-vis des erreurs d'estimation de ces paramètres est étudiée.

Soit f le facteur correctif lorsque les paramètres sont connus ($\varphi_A, R, h_{\text{ambiguïté}}$). En pratique, ces paramètres ne sont pas connus et doivent être estimés. Soit \hat{f} le facteur correctif pour les paramètres estimés ($\hat{\varphi}_A, \hat{R}, \hat{h}_{\text{ambiguïté}}$).

2.5.3.2.1. Erreur sur $\hat{\varphi}_A$

Pour déterminer l'influence de $\hat{\varphi}_A$ sur le facteur correctif, le contraste R est supposé connu. Pour une erreur d'estimation de la hauteur apparente de $\pm 85\%$, les facteurs correctifs sont représentés en rouge sur la figure 55a. La courbe noire en traits pleins correspond au facteur correctif lorsque les paramètres sont connus. Lorsque $\hat{\varphi}_C$ est sur estimé (ou $\hat{\varphi}_A$ est sous estimé) de 14m pour une hauteur apparente initiale de 16m, l'erreur sur le facteur correctif interférométrique est faible. Le modèle proposé est donc robuste vis-à-vis des erreurs d'estimation sur les phases. Cette erreur reste très faible comparée aux erreurs potentielles dues à un moyennage spatialement restreint.

2.5.3.2.2. Erreur sur \hat{R}

De même, l'influence d'une erreur d'estimation de R sur le facteur correctif est représentée sur la figure 55b. En traits continus rouges sont représentés les facteurs correctifs interférométriques lorsque R est mal estimé de $\pm 30\%$. Pour une sous estimation du contraste de 3dB, le facteur correctif varie peu également. Une fois encore, l'influence de l'erreur d'estimation sur R reste faible par rapport à celle provenant du moyennage spatial.

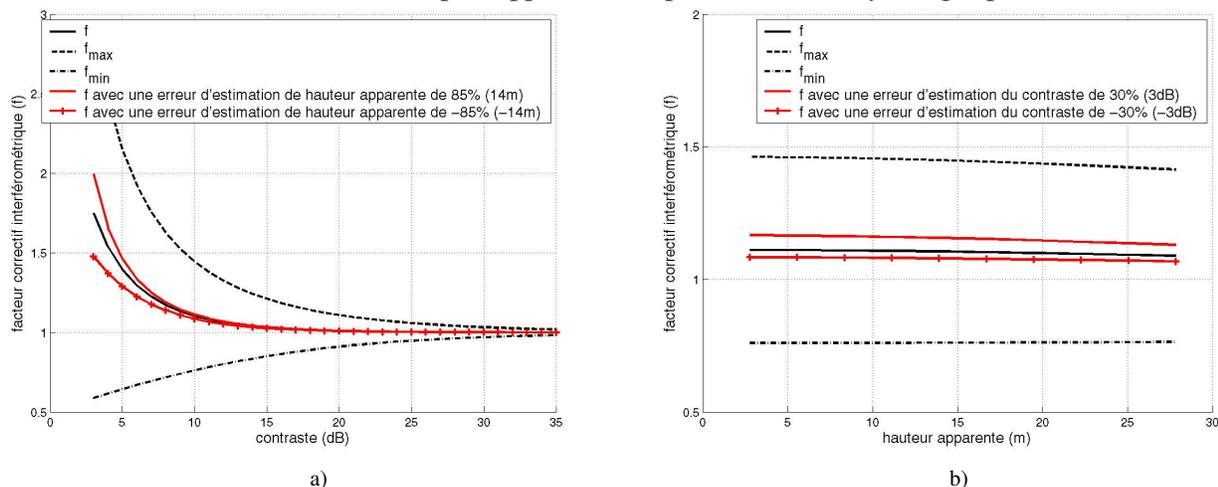


Figure 55 – Influence de a) $\hat{\varphi}_A$ et $\hat{\varphi}_C$ et b) de \hat{R} sur le facteur correctif

2.5.4. Validation du modèle

Ce modèle est validé sur l'extraction de façades de bâtiments. Les sommets des façades des bâtiments remplissent la condition requise d'un mélange de deux contributions d'altitudes

différentes en début de zone de repliement (cf. figure 56). Le repliement éventuel entre l'écho double et le toit du bâtiment est supposé nul car l'écho double apparaît généralement avec une amplitude très élevée.

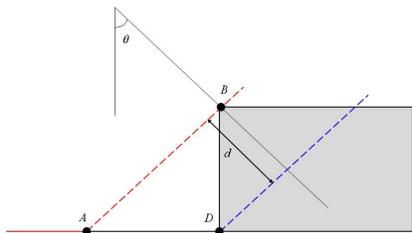


Figure 56 –Repliement à deux contributions d'altitudes différentes sur la façade d'un bâtiment

2.5.4.1. Méthode de validation

Pour chaque façade, deux mesures sont à disposition :

- la mesure monoscopique (h_{mono}) entre les deux limites de la façade
- la mesure de hauteur interférométrique sans utilisation du modèle ($h_{apparente}$)

On détermine la mesure de hauteur interférométrique avec utilisation du modèle (h_{modele}).

Ces trois hauteurs s'écrivent:

$$h_{mono} = \frac{d}{\cos(\theta)} \quad h_{apparente} = \frac{h_{amb}(\varphi_C - \varphi_A)}{2\pi} \quad h_{modele} = \frac{h_{amb}(\varphi_B - \varphi_A)}{2\pi}$$

où h_{amb} est la hauteur d'ambiguïté

Pour estimer le paramètre φ_A (phase au sol), une croissance de région est utilisée. Le point d'accès est situé à une distance légèrement plus faible que le bord élevé de la façade. Pour diminuer le bruit sur l'interférogramme, cette croissance de région est effectuée sur un voisinage 8x8.

Pour estimer A, les profils d'amplitude et de phase moyenne du segment représentant le bord élevé de la façade sont calculés pour des distances décroissantes. Le segment présentant la phase la plus proche de celle de la croissance de région est retenu. A est choisi comme l'amplitude de ce segment. L'amplitude de la zone obtenue par croissance de région n'est pas utilisée car elle peut contenir des points brillants qui fausseraient l'estimation de A.

Pour valider le modèle, 18 façades, présentant des échos de bord de toit ou des textures avec des contrastes allant de 5 à 25dB environ, sont considérées. Les extrémités de chaque façade sont extraites manuellement. C et φ_C sont ainsi obtenus. Quelques exemples sont fournis sur la figure 57. En bleu, sont représentés les échos sélectionnés. En rouge, sont délimitées les zones à partir desquelles les croissances de région ont été effectuées.

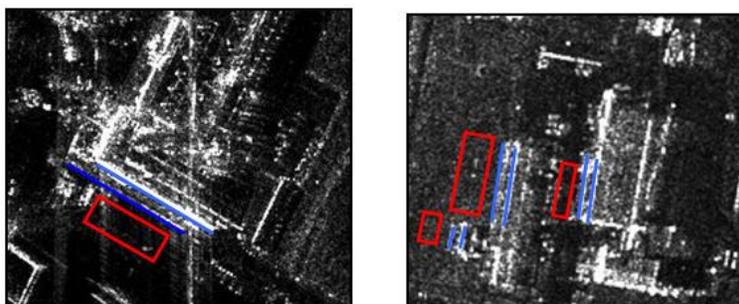


Figure 57 – Exemples de façade extraite

Le critère d'évaluation du modèle consiste en l'adéquation des mesures monoscopiques et interférométriques (avec ou sans modèle).

2.5.4.2. Altitude de référence

L'altitude de référence est l'altitude de la base de la façade. Deux choix sont possibles. L'altitude de référence est soit l'altitude de la zone extraite par croissance de région, soit l'altitude de l'écho double. Dans le premier cas, il a été préalablement vérifié que le repliement des façades avait effectivement lieu sur le sol. Dans le second cas, l'altitude de l'écho double est obtenue lors de l'extraction manuelle de la façade.

2.5.4.3. Résultats

Sur les deux graphiques de la figure 58 sont représentés en noir et en rouge les rapports des mesures interférométriques, respectivement sans et avec modèle, et monoscopiques en fonction du contraste. Les régressions linéaires sur ces mesures sont également présentées (en noir sans modèle et en rouge avec modèle). Sur les graphiques a et b, les altitudes de référence sont respectivement le sol et l'écho double.

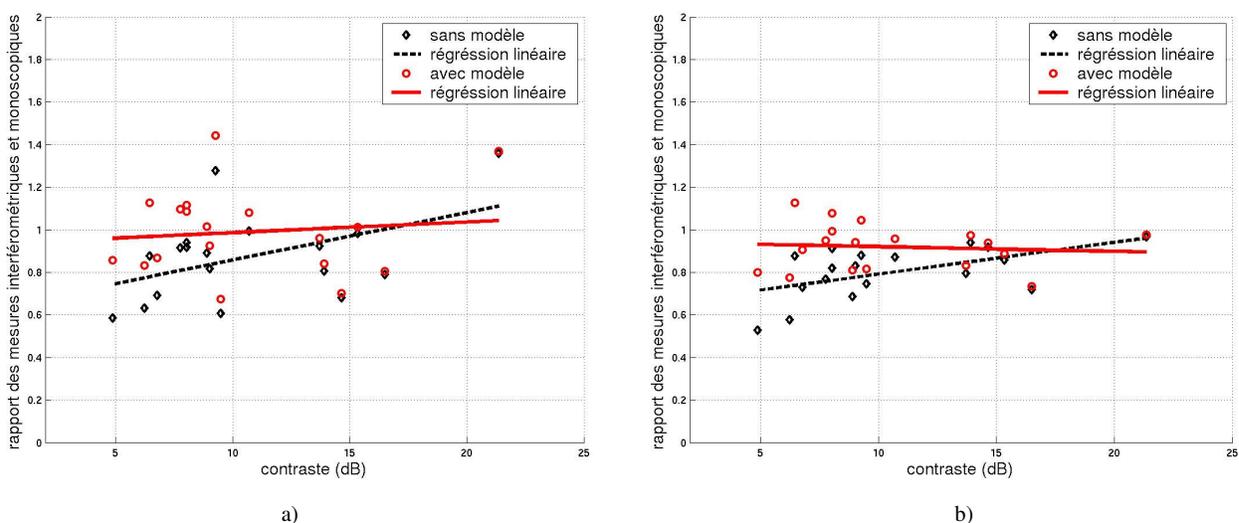


Figure 58 – Validation du modèle avec l'altitude de a) la croissance de région et b) l'écho double comme référence

En utilisant l'altitude du sol devant la zone de repliement comme référence (figure 58a), la moyenne du rapport des mesures interférométriques sans modèle et des mesures monoscopiques est de 0.870 pour un écart type de 0.207. La régression linéaire, en traits discontinus noirs, montre que ce rapport dépend fortement du contraste.

Le modèle proposé permet d'améliorer l'estimation de la hauteur interférométrique de la façade. La moyenne du critère vaut alors 0.989. En revanche, l'écart type reste important (0.204). Ceci est dû à l'écart type altimétrique important de la mesure interférométrique sur le sol pour des rétrodiffusions de type Rayleigh (cf. §2.3.2). De plus, l'attitude de l'avion peut varier suffisamment pour que l'altitude interférométrique du sol ne soit pas constante sur une zone. Néanmoins, la régression linéaire est quasi horizontale et traduit une faible dépendance vis-à-vis du contraste.

Lorsque l'écho double sert de référence (figure 58b), la mesure d'altitude de référence est plus localisée et correspond à l'altitude au pied du bâtiment, ce qui se traduit par un écart type sur le critère plus faible (0.121 sans modèle et 0.108 avec modèle). Il est également visible que la droite correspondant à la régression des rapports est une fois de plus quasiment horizontale. Ce qui signifie que l'estimation de hauteur des façades peut être effectuée

indépendamment du contraste. Il est constaté au §2.4.4.2 que l'altitude de l'écho double est légèrement supérieure à celle du sol (sur les bâtiments testés, un biais de 50cm est présent). Ce biais provoque une sous estimation de la hauteur des façades. En effet, la moyenne du critère est plus faible (0.800 sans modèle et 0.919 avec modèle).

L'erreur quadratique moyenne sur l'estimation de la hauteur interférométrique des façades est de 2.24m sans modèle et de 1.26m avec modèle.

2.5.5. Limites du modèle

Ce modèle est utilisable seulement si le repliement a lieu sur une zone dont il est possible de déterminer l'altitude. Sur l'exemple de la figure 19b en p39, la façade du bâtiment se replie sur de l'ombre. La mesure interférométrique de la façade correspond alors directement à son altitude réelle. Une séparation de mélange dans ce cas fournit des résultats aberrants. Il faut également que la hauteur de chacune des contributions soit constante sur la zone où est effectuée le moyennage, ce qui exclut d'utiliser ce modèle sur certains toits en pente.

2.5.6. Conclusion

Un modèle visant à séparer les contributions d'un mélange interférométrique de deux zones d'altitudes différentes a été proposé. Ce modèle permet, après moyennage spatial, de séparer ces deux contributions indépendamment du contraste et d'améliorer l'estimation de la hauteur interférométrique des façades de bâtiments. Ce modèle est à rapprocher des algorithmes de séparation de mélanges en polarimétrie interférométrie (cf. §6.2).

2.6. Conclusion

Des exemples de signatures de bâtiments sur images RAMSES haute résolution ont été présentées. Ces signatures peuvent se composer d'échos forts (simples ou multiples), de textures et d'ombres. L'origine de ces primitives a été expliquée par une approche géométrique et/ou phénoménologique. Cette analyse permet de soulever l'une des difficultés rencontrées par le traitement automatique des scènes urbaines sur images RSO : le caractère imprévisible de la présence des primitives constituant la signature RADAR des bâtiments. Par exemple, des textures apparaissent sur les toits des bâtiments seulement s'ils sont rugueux à l'échelle de la longueur d'onde. Les échos du bord du toit ne sont visualisables que si un élément sur haut de la façade (gouttière...), réfléchit le faisceau électromagnétique vers le capteur. Ces primitives permettent d'obtenir des informations 3D grâce aux mesures monoscopiques et interférométriques. Les mesures monoscopiques, lorsqu'elles sont valides, s'avèrent plus précises que les mesures interférométriques. Enfin, un modèle de mélange altimétrique de deux zones d'altitudes différentes a été proposé et validé. La séparation de ce mélange a permis d'améliorer l'estimation de la hauteur interférométrique des façades.

Toutes les mesures de 3D détaillées ont un cadre d'application restreint. Le repliement affecte les mesures monoscopiques et les mesures interférométriques. La mesure monoscopique sur la parallaxe n'est applicable que sur des primitives appartenant à la même façade. Néanmoins, ces mesures peuvent être utilisées de façon complémentaire. Par exemple la détection de façade à partir d'une ombre ou de deux échos utilise conjointement la mesure interférométrique et la mesure monoscopique. Si la façade est détectée en utilisant la mesure sur la parallaxe, il est possible d'en déduire la zone où apparaît le repliement et où

l'information interférométrique n'est pas utilisable. Si la façade n'est pas détectée, cela signifie que les échos n'appartiennent pas à la même façade. A partir d'une ombre, si la façade n'est pas détectée, cela implique que l'ombre a été tronquée et qu'il y a du repliement à une des extrémités de l'ombre. Une extraction fiable du 3D sur un couple d'images interférométrique RSO ne peut donc se faire uniquement qu'en utilisant conjointement toutes les mesures disponibles.

3. Etude bibliographique sur la représentation 3D des bâtiments en imagerie RSO

3.1. Introduction

Pour représenter une scène en trois dimensions (3D), plusieurs approches sont envisageables. La première utilise un modèle numérique de surface (MNS) où aucune hypothèse sur la scène n'est effectuée. Le sol et les zones surélevées de la scène sont traités indifféremment. Dans un MNS l'information 3D est locale. Sur l'exemple de la figure 59a, le MNS d'un bâtiment à toit plat sur sol plat est représenté. L'information altimétrique sur le sol et le toit n'est pas constante à cause du bruit sur les mesures.

La seconde approche est également surfacique mais effectue une distinction entre sol et zone élevée. La première étape consiste à déterminer la hauteur du sol pour obtenir le modèle numérique de terrain (MNT). Le sursol est ensuite extrait par un traitement différent.

Enfin la dernière approche utilise une représentation sous forme d'objets des éléments de la scène pour former un modèle numérique d'élévation (MNE). Des hypothèses sur les objets représentés sont alors nécessaires. Par exemple, le sol de la scène peut être représenté par un sol plat. Les zones élevées peuvent être représentées par des parallélépipèdes rectangles (cf. figure 59b). Elles permettent d'obtenir un ensemble d'objets types, appelés modèles, caractérisés par des paramètres géométriques. Cette étape est encore appelée reconstruction.

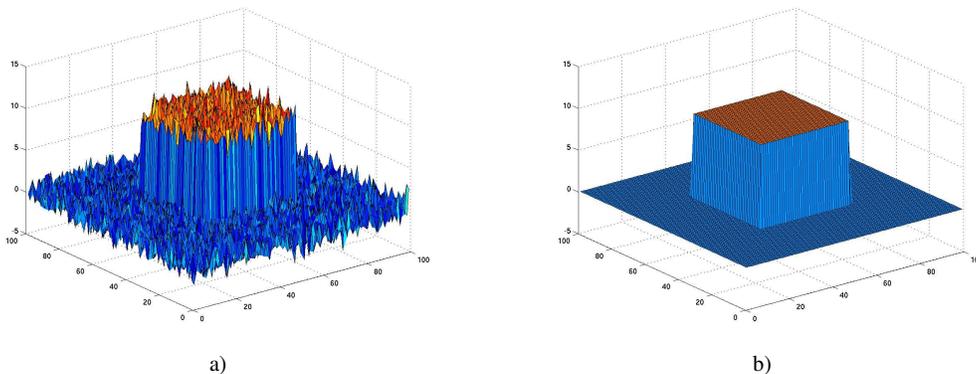


Figure 59 – a) MNS d'une zone avec un bâtiment et b) MNE correspondant

Dans ce chapitre, l'état de l'art sur la représentation 3D de bâtiments à partir d'images RSO est effectué. Pour cela, les données utilisées sont tout d'abord présentées. Puis, un résumé des prétraitements utilisés est effectué. Les algorithmes d'extraction du 3D sont détaillés et regroupés selon l'approche surfacique (MNS ou MNT/zone élevées) ou objet (MNE).

3.2. Présentation des données

Dans ce paragraphe le type de données utilisées pour la reconstruction de bâtiments est présenté. Les algorithmes considérés utilisent des images hautes résolutions (de 20cm à 2-3m environ) et moyennes résolutions (2-3m à 10m environ).

[Gamba 1999, 2000a, 2000b] : les données utilisées sont issues du capteur AIRSAR (NASA/JPL) monté sur un DC8. Elles sont traitées par le système TOPSAR. Ils travaillent en bande C et la précision nominale du capteur en hauteur est 2.5m. Les images originales ont une résolution en distance de 3.75m qui passe à 7.5m après le déroulement de phase (moyennage de deux pixels). La résolution en azimut est également 7.5m. Les auteurs travaillent sur des images de Santa Monica : il y a de grands bâtiments (entre 40m et 110m de hauteur). Les traitements sont effectués en géométrie RADAR.

[Hoepfner 1999], [Bolter 2001] : le capteur est le Sandia Spotlight (aéroporté). Les images sont prises en bande Ku et ont une résolution de 33cm. 4 images de la même zone (site de MOUT), où il y a une dizaine de bâtiments, d'empreintes rectangulaires, avec des toits de structures diverses (plats, doubles pentes...), ont été prises avec des trajectoires orthogonales. Les bâtiments sont proches et le repliement tronque certaines ombres. L'angle d'incidence est de 45°. La dynamique des images en amplitude est réduite (14 niveaux de gris, quoique le capteur ait 32 bits de dynamique).

[Quartulli 2001, 2003a, 2003b, 2004] : le capteur utilisé est celui d'Intermap Technologies avec une visée de 45°. Les images sont prises en bande X avec une résolution de 2m (images 8 vues). Les images sont acquises sur le site industriel de Oberpfaffenhofen. Les bâtiments ont des empreintes rectangulaires et sont entourés de routes.

[Dell'Acqua 2001], [Michaelsen 2002], [Stilla 2001, 2003a, 2003b] : le capteur utilisé est AeR-II du FGAN. Les images sont en bande X (10GHz, longueur d'onde 3cm), avec une résolution de 1m environ, l'écart-type altimétrique, mesuré sur une zone plate, est au mieux de 2m. Les images portent sur le site de l'aéroport de Francfort. Il y a de grands bâtiments proches avec des formes complexes.

[Simonetto 2002] : les données utilisées sont issues du capteur RAMSES de l'ONERA en bande S et X. Les images sont acquises sur la zone industrielle du port de Dunkerque et la résolution des images est de l'ordre du mètre. Les algorithmes sont utilisés sur image géoréférencées.

[Sorgel 2003a] : le capteur utilisé est DO-SAR. Les images sont acquises en bande X (9,5GHz, longueur d'onde 3cm) et la résolution en géométrie RADAR est de 1.2m. Les auteurs disposent de plusieurs acquisitions interférométriques (amplitude / interférogramme / cohérence) effectuées avec des trajectoires orthogonales. La zone est rurale (banlieue) et peu dense. Les algorithmes sont utilisés sur les données en géométrie RADAR pour éviter les distorsions dues au géoréférencement. La scène contient plusieurs bâtiments de type différents (en un seul bloc ou avec extensions). A partir d'une étendue large de terrain, l'écart-type altimétrique sur l'interférogramme est estimé à 1m environ.

[Petit 2004], [Tison 2004], les données utilisées sont les images RAMSES (Bande X) et d'autres issues du capteur AeS-1. Les images retenues pour les tests sont celles de la cité universitaire, du CNES, du Grand Rond et de la prison Saint Michel à Toulouse. Les résolutions en distance et en azimut des images aerosensing sont respectivement 37cm et 18cm. Les images RAMSES proviennent de la même campagne que celles présentées en annexe A et ont les mêmes résolutions (50cm environ).

[Thiele 2006] : l'auteur utilise des données issues du capteur d'Intermap Technologies, à des résolutions de 38cm en distance et 16cm en azimut. Elle dispose de deux traces orthogonales de la même zone. L'acquisition est effectuée en mode ping-pong : les antennes maître et esclave émettent tour à tour l'onde RADAR. Cette technique permet de simuler une

base interférométrique deux fois plus grande. La précision interférométrique est donc améliorée par rapport à une acquisition traditionnelle. Néanmoins cette précision n'est pas indiquée.

[Michaelsen 2005], [Soergel 2006] : deux images d'un même bâtiment, à des résolutions de 20cm et 40cm, issues de trajectoires orthogonales sont utilisées. Elles proviennent du capteur aéroporté PAMIR (FGAN) en bande X. Le bâtiment étudié est très complexe. Une analyse de la signature de ce bâtiment est effectuée sur l'image à 20cm dans [Stilla 2005]. Les traitements sont effectués sur images géoréférencées en géométrie terrain. Il n'y a pas de données interférométriques.

Plusieurs auteurs utilisent deux voire quatre acquisitions de la même scène avec des trajectoires orthogonales. L'intérêt d'une telle approche a été montré théoriquement dans [Bolter 2001]. L'auteur constate que, sur image simulée, la reconstruction est de meilleure qualité pour un angle d'incidence de 60°. Cette approche est confirmée dans [Soergel 2003b]. Les auteurs utilisent un simulateur sur le modèle numérique du campus universitaire de Karlsruhe afin de déterminer quantitativement l'apport de nouvelles traces provenant des directions Est, Ouest, Sud et Nord. Les images simulées sont comparées aux images réelles prises en bande X (résolution de 1m). Pour chaque trace les images ont été simulées pour un angle de prise de vue entre 30° et 70°. L'angle permettant une visibilité optimale des bâtiments sur chaque image a été déterminé à 60° pour les traces Est, Ouest, Sud et 70° pour la trace Nord. Pour les routes, la visibilité est maximale pour un angle de visée de 45° sur les directions Nord et Ouest qui correspondent aux directions privilégiées du réseau. Les résultats sont consignés dans le tableau ci-dessous. L'utilisation de plusieurs vues apporte un supplément d'information non négligeable. Utiliser deux et quatre vues permet d'augmenter la surface visible des toits respectivement de 21% et 35%.

Traces RSO	1	2	3	4
Surface de toit visible en %	51.6	72.8	81.5	86.5
Surface de route visible en %	39	47	55.6	62

L'utilisation de plusieurs vues permet également d'améliorer la localisation des bâtiments car plus d'information sur les façades sont disponibles grâce aux ombres [Bolter 2001] et aux échos forts [Thiele 2006], [Soergel 2006].

3.3. Prétraitement

La présence de chatoiement sur les images RADAR rend les traitements difficiles. C'est pourquoi un prétraitement s'avère souvent nécessaire. Il consiste en :

- un passage en multivue ([Gamba 2000a], [Simonetto 2002], [Tison 2004], [Quartulli 2004]) pour réduire l'aspect « granuleux » au prix d'une perte de résolution.
- le filtre de Desnos et Matteini, [Desnos 1993] ([Michaelson 2002], [Sorgel 2003a]). Ce filtre prend en compte la nature statistique et géométrique de l'image en combinant le filtre de Lee amélioré (cf. §4.2.2.1), un détecteur de contour et un détecteur de ligne. Dans les zones où ne sont détectés ni contours ni lignes, le filtre de Lee amélioré est appliqué. Si un bord ou une ligne est détecté, le filtre de Lee est utilisé sur une zone excluant respectivement une moitié ou la bande centrale de la fenêtre de filtrage.
- un filtre médian sur les images en amplitude et sur l'interférogramme ([Dell'Acqua 2001], [Bolter 2001])

- séparation des points brillants à plus de 40dB pour effectuer un étalement non linéaire (non détaillé) de l'image sur 1 octet [Soergel 2006].

3.4. Approches surfaciques

3.4.1. Extraction du modèle numérique de surface

L'extraction du modèle numérique de surface peut s'effectuer à partir des données interférométriques à savoir l'image d'amplitude et l'interférogramme (approche mixte) ou des données radargrammétriques. La radargrammétrie reprend le principe de la stéréoscopie en imagerie optique. L'information 3D est obtenue à partir de la différence de localisation sur les deux images du couple radargrammétrique d'un même point dans l'espace. Pour cela les images du couple doivent être acquises à des angles d'incidence très différents.

3.4.1.1. Approche mixte

L'interférogramme est une carte dense de phase. Il permet d'obtenir une information altimétrique relative en chaque point de la scène. Pour obtenir les hauteurs absolues sur la scène il faut connaître l'altitude d'un point de référence. Néanmoins, l'interférogramme ne peut pas être considéré directement comme un MNS. En effet, l'interférogramme doit être géoréférencé, c'est-à-dire projeté dans le repère monde. De cette façon les distorsions géométriques provenant de l'acquisition en distance sont éliminées. De plus, dans les zones d'ombre, la phase de l'interférogramme correspond à du bruit. Elles sont donc extraites puis masquées pour ne pas être prise en compte. Les algorithmes d'extractions des ombres sont présentés au §3.5.3.1.

Les algorithmes présentés utilisent la classification Markovienne dont le principe est détaillé en annexe C. Dans [Frery 1999a,1999b], la distribution inverse généralisée d'une gaussienne est utilisée pour modéliser les lois présentes sur les images (cf. annexe B.3). Pour effectuer une classification par champ de Markov, les auteurs utilisent un algorithme ICM-ML. L'état initial est déterminé par le critère du maximum de vraisemblance (Maximum Likelihood). Une optimisation ICM (Iterative Conditional Mode) est ensuite utilisée.

Dans [Tison 2003], l'auteur utilise la distribution de Fisher (Pearson VI) pour modéliser les statistiques de l'image en amplitude. La classification Markovienne sur l'image d'amplitude est basée sur les logcumulants [Nicolas 2003]. 6 classes sont considérées, à savoir, le sol, la végétation sombre, la végétation claire, les toits sombres, les toits moyens et les toits clairs/forts réflecteurs [Tison 2003], [Tison 2004]. Une régularisation est effectuée sur l'interférogramme pour réduire le bruit interférométrique. D'autres détecteurs (routes, ombres, coins réflecteurs) sont utilisés. A partir de ces informations, l'auteur effectue une nouvelle classification avec optimisation conjointe de la classe et de la hauteur en utilisant des informations contextuelles entre les classes.

3.4.1.2. Approche radargrammétrique

L'intérêt de la radargrammétrie par rapport à l'interférométrie est de fournir une information altimétrique absolue et non relative. Pour obtenir un MNS à partir d'un couple radargrammétrique, un appariement dense des deux images est nécessaire. Les images du couple radargrammétrique sont prises avec un grand écart entre les angles d'incidences. Par conséquent, les distorsions géométriques et la phénoménologie RADAR sont très différentes d'une image à l'autre. Par exemple, le critère de Rayleigh (cf. §2.4.3) tient compte de l'angle d'incidence pour caractériser la rugosité d'une surface. Sur les deux images d'un couple

radargrammétrique, un même toit peut ne pas apparaître avec les mêmes textures. La mise en correspondance dense sur toute l'image est donc délicate.

Plusieurs critères de corrélations provenant des deux approches ont été testés dans [Simonetto 2002], à savoir :

- corrélation sur les différences
- corrélation sur les covariances
- corrélation de rang
- corrélation sur les différences sur le logarithme des images
- corrélation sur les covariances sur le logarithme des images
- corrélation adaptée à la loi Gamma [Nocera 1996].

Les trois premiers corrélateurs sont classiques à la mise en correspondance sur image optique. Par le passage au logarithme, l'auteur tente de se ramener au cas optique. En effet, l'étude statistique de l'annexe B.3 montre que la loi lognormale pourrait être utilisée pour estimer les statistiques des images RADAR. Le passage au logarithme des images permet alors de se ramener au cas gaussien. Le dernier corrélateur est adapté à la loi Gamma. Les deux premiers critères obtiennent les meilleures performances. L'étude de ces deux corrélateurs a également permis de déterminer un rayon minimal de la fenêtre du corrélateur de 5 pixels pour avoir une bonne mise en correspondance et de définir un seuil pour conserver entre 70 et 80% des points appariés. L'auteur conclut également sur le fait qu'un critère sur la hauteur relative du pic principal par rapport aux secondaires ou sur la largeur du pic principal n'est pas envisageable car les courbes de performances sont trop bruitées.

Lorsque la mise en correspondance est effectuée, une carte de disparité est calculée. La disparité en un point de la scène est la distance entre les projetés de ce point sur les deux images du couple radargrammétrique. Cette disparité est directement reliée à l'altitude de la scène.

3.4.2. Extraction du modèle numérique de terrain

Le modèle numérique de terrain est extrait sur données interférométriques ou radargrammétriques.

3.4.2.1. Approche interférométrique

Le MNT peut être obtenu par filtrage de l'interférogramme. La taille de la fenêtre est idéalement choisie plus grande que le bâtiment le plus important. Dans [Bolter 2001], un filtre minimum est utilisé. Ce filtre permet d'obtenir la valeur minimum de la fenêtre de filtrage : il est donc très sensible au bruit interférométrique. C'est pourquoi l'interférogramme est filtré, dans [Soergel 2003a], par le filtre développé par Eckstein et Munkelt ([Eckstein 1995]). Il consiste à utiliser consécutivement deux filtres de rang (r et $1-r$) qui donnent les $r^{\text{ième}}$ et $N-r^{\text{ième}}$ valeurs classées par ordre croissant où N est le nombre de pixels de la fenêtre choisie. Le premier filtre (de rang r) permet d'obtenir un MNT bruité. A l'issue de ce premier filtrage, un lissage est opéré par le second filtre. Pour diminuer le bruit sur l'interférogramme, les pixels présentant une élévation importante mais une amplitude ou cohérence faible sont rejetés. A partir d'une zone de sol à hauteur constante sur l'interférogramme, les auteurs calculent l'écart-type sur la mesure altimétrique et obtiennent une valeur de seuil pour extraire le MNT et les zones élevées.

Dans [Petit 2004], la logique floue est utilisée pour distinguer le terrain et produire un masque urbain. Un premier masque grossier du sol est obtenu à partir de classifications floues pixeliques effectuées sur l'image d'amplitude (ombre, goudron, herbe, végétation, structures, points brillants, points très brillants), l'interférogramme (bas, haut) et sur la cohérence

(cohérence proche de 1 ou non). Le masque de sol « grossier » est formé par la réunion des pixels des classes sol, herbe et goudron qui sont fortement cohérent.

La hauteur moyenne du masque « grossier », notée h_G , est alors calculée. Une estimation de la hauteur moyenne du sol est obtenue en calculant la moyenne des pixels du masque de sol « grossier » ayant une hauteur inférieure à h_G .

3.4.2.2. Approche radargrammétrique

Dans [Simonetto 2002], le MNT est déterminé par la carte de disparité dense en multivue sur les zones ne correspondant pas aux échos forts et aux bâtiments. Une interpolation est alors effectuée pour déterminer l'altitude sur les zones masquées par les bâtiments.

3.4.3. Extraction des zones élevées

Le sursol peut être extrait en utilisant uniquement l'amplitude, uniquement l'interférogramme, ou en utilisant conjointement ces deux images (approche mixte).

3.4.3.1. Approche amplitude

Les champs de Markov sont utilisés dans [Kirscht 1998] pour classifier les images d'amplitude en trois classes (sol, ombres et zones élevées). La distinction entre classe est effectuée sur les moments du premier et second ordre des statistiques associées aux textures. Le choix des statistiques n'est pas précisé. La distinction sol / sur-sol est effectuée en considérant un modèle de terrain de la zone. Les hauteurs des bâtiments sont estimées par mesure monoscopique.

3.4.3.2. Approche interférométrique

Dans [Gamba 1999, 2000a, 2000b], après un passage en multivue, les auteurs utilisent l'algorithme développé par Jiang et Bunke pour détecter les bâtiments sur l'interférogramme ([Jiang 1994]). Cette technique s'appelle la 'scan line approximation' : en considérant une ligne ou une colonne de l'image dans l'espace 3D (la troisième dimension correspondant à la valeur des pixels), une courbe, divisible en plusieurs sections ayant chacune des paramètres différents, est obtenue. Ces morceaux de courbes sont approchés itérativement par des segments linéaires. La jonction entre chaque segment est un pixel qui pourrait appartenir au contour. Une mesure de continuité est alors employée pour déterminer si effectivement ce pixel appartient au contour. Sur chaque 'scan line' (ligne ou colonne) une série de segments est donc obtenue. Trois segments similaires, de longueur supérieure à un seuil, et appartenant à des lignes différentes, sont nécessaires pour être un initiateur de la croissance de régions qui s'effectue non pas sur les pixels mais sur les segments. Si un segment est suffisamment proche (seuil à fixer) du plan formé par les segments de la région à l'itération précédente, il est ajouté à la zone. La hauteur du bâtiment est obtenue par l'interférogramme sur les zones segmentées.

Dans [Bolter 2001], l'auteur dispose de quatre traces orthogonales du site de MOUT et fusionne les interférogrammes par le maximum des hauteurs. Les bâtiments sont correctement extraits mais la précision planimétrique est mauvaise à cause du repliement. L'auteur a également comparé cet algorithme et celui des 'scans lines' et conclut que ce dernier aboutit à des bâtiments moins bien extraits et qu'il y a des fausses alarmes.

Dans [Soergel 2003a], après avoir filtré l'interférogramme, les zones élevées sont extraites par seuillage sur l'interférogramme (cf. §3.4.2.1). Sur chaque zone élevée, l'élévation, pondérée par la cohérence, est moyennée. La hauteur estimée s'exprime par :

$$h_{\text{estimée}} = \frac{\sum_p \gamma_p \cdot h_p}{\sum_p \gamma_p} \quad (65)$$

où p désigne les pixels de la région sur laquelle s'effectue l'estimation. h_p et γ_p sont respectivement la hauteur et la cohérence associées à ces pixels. Les pixels les plus cohérents, et dont la hauteur interférométrique est la moins bruitée, sont prépondérants dans le calcul de la hauteur estimée.

3.4.3.3. Approche mixte

Dans [Soergel 2000], [Dell'Acqua 2001], [Stilla 2001, 2003a] et [Michaelsen 2002], après avoir débruité l'image d'amplitude [Desnos 1993], le gradient de l'image est calculé. Les pixels avec un fort gradient sont caractéristiques des contours entre deux régions. Un seuillage sur l'image de gradient permet alors d'obtenir des régions. Ce seuil est fixé à la moyenne plus deux fois l'écart-type de l'image en amplitude. Puis une croissance de région sur les images en amplitude est effectuée avec comme contrainte l'interdiction de traverser les frontières déterminées par l'étape précédente. Pour éviter la sous segmentation, les paramètres de la croissance de régions sont réglés pour obtenir un maximum de régions.

Pour chaque région issue de la segmentation, la hauteur est estimée en pondérant et normalisant l'interférogramme par l'intensité :

$$h_{\text{estimée}} = \frac{\sum_p i_p \cdot h_p}{\sum_p i_p}$$

où p désigne les pixels de la région segmentée sur laquelle s'effectue l'estimation et i_p l'intensité du pixel p . Utiliser l'intensité plutôt que la cohérence pour la pondération permet de privilégier les forts réflecteurs. En effet, en imagerie RSO aéroporté monopasse, la cohérence moyenne est très élevée et, en utilisant (65), le sol et les forts réflecteurs sont peu différenciés.

Les régions présentant une intensité et une cohérence faibles sont considérées comme des zones d'ombre ou de route. Cette hypothèse n'est pas valide dans le cas de toits non-rétrodiffusants. Pour être validés en tant que bâtiment, les zones élevées doivent avoir une de ces régions à proximité. Ensuite, une analyse des ombres est menée : si l'ombre projetée par le bâtiment présente deux régions distinctes alors cela signifie que le bâtiment est sous segmenté et doit être séparé en deux. Si des régions contiguës présentent des hauteurs moyennes proches, alors elles sont fusionnées pour éviter la sur segmentation. Aucune indication n'est fournie sur les seuils nécessaires à cet algorithme de fusion.

3.4.4. Conclusion

Dans ce paragraphe les différentes techniques d'extraction des MNS, MNT et zones élevées ont été détaillées. La radargrammétrie peut être utilisée pour obtenir le MNT et, plus difficilement le MNS. Les extractions des zones élevées peuvent être effectuées sur l'image d'amplitude, sur l'interférogramme ou grâce à l'utilisation conjointe de ces deux images. L'interférogramme est l'image la plus utilisée pour l'extraction des zones élevées car elle permet d'accéder à une information altimétrique dense.

3.5. Approche objet : extraction du modèle numérique d'élévation

3.5.1. Introduction

La reconstruction est l'étape permettant d'obtenir le MNE, c'est-à-dire une représentation sous forme d'objets de la scène. Pour cela la notion de modèles de bâtiments est introduite pour décrire les bâtiments à reconstruire. Des hypothèses sur le type de bâtiments sont alors nécessaires pour définir les modèles. La reconstruction peut s'effectuer par une méthode ascendante (« image-based ») ou descendante (« model-based »).

La méthode ascendante utilise les informations bas niveau de l'image pour aboutir au modèle (objet haut niveau). Elle peut être utilisée sur les zones élevées (cf. §3.4.3) et/ou à partir des primitives (ombres, échos...). Dans le deuxième cas, les primitives sont alors assemblées pour former des bâtiments en accord avec les modèles choisis. La méthode descendante considère chaque modèle et trouve la configuration optimale, selon un critère à définir, permettant d'aboutir à la signature du bâtiment.

Dans un premier temps les modèles utilisés dans la littérature sont décrits. Puis les différents algorithmes d'extraction de primitives sont présentés. Ensuite l'utilisation des primitives extraites pour obtenir une information de hauteur est détaillée. Enfin, les algorithmes de reconstruction sont présentés.

3.5.2. Modélisation des bâtiments

La modélisation d'un bâtiment se présente sous la forme d'un ensemble de contraintes géométriques que le bâtiment doit suivre. Voici une liste des règles couramment utilisées ([Bolter 2001], [Dell'Acqua 2001], [Stilla 2001], [Simonetto 2002], [Soergel 2003a, 2006], [Quartulli 2001, 2003a, 2004], [Petit 2004], [Thiele 2006]) ; les bâtiments sont modélisés comme des structures :

- avec une empreinte rectangulaire ou de forme plus complexe à angles droits ou en forme de cylindres (caves)
- ayant des toits plats, en pente ou à double pente
- suivant une orientation globale

Lorsque l'empreinte est à angle droit [Soergel 2003a], cela revient à modéliser les bâtiments par un ensemble de briques élémentaires d'empreintes rectangulaires.

Une autre approche est choisie dans [Quartulli 2001, 2003a, 2004] où chaque bâtiment est décomposé en un ensemble de facettes modélisées par leurs centroïdes, largeurs, hauteurs et réflectivité RADAR.

Ces hypothèses sont suffisamment générales pour permettre de décrire un grand nombre de bâtiments. Néanmoins, les bâtiments, sortant du cadre fixé par ces hypothèses, ne pourront pas être reconstruits correctement.

3.5.3. Extraction des primitives

3.5.3.1. Extraction des ombres

Pour extraire les ombres, dans [Hoepfner 1999] et [Bolter 2001], un filtre binaire rotatif, suivant plusieurs directions (36 pour Hoepfner et 16 pour Bolter) est utilisé sur l'image d'amplitude.

Hoepfner s'intéresse à la détection du contour entre le bâtiment et l'ombre, encore appelé « backedge ». Pour cela, il utilise un critère de force sur l'orientation du contour associé à chacune des 36 directions. Lorsque la direction du contour de l'image et la direction du filtre sont les mêmes, le critère calculé est élevé. Un seuil ad-hoc sur ce critère est utilisé pour sélectionner les pixels appartenant au contour de l'ombre. Pour éviter les fausses alarmes (extraction d'un écho double par exemple), une fenêtre glissante est utilisée pour vérifier qu'il y a une zone de faible amplitude à une distance légèrement plus éloignée du contour.

Bolter utilise ce filtre pour extraire complètement les ombres et pas seulement le backedge. Le critère de force utilisé dépend de la différence de moyenne et de variance des zones situées de part et d'autre du contour hypothétique. Lorsque l'orientation du filtre et du contour du bâtiment sont les mêmes, et lorsque les zones de part et d'autre du centre de la fenêtre de filtrage sont contrastées, la force de ce contour est importante. Une fois encore, des seuils, peu intuitifs et peu adaptés à des images de forte dynamique sont utilisés. Cet algorithme ne semble donc pas adapté aux images RAMSES.

Dans [Soergel 2003a], les auteurs extraient les ombres par la fusion de deux sets. Le premier provient du détecteur de contour Canny (généralement ceux entre le toit du bâtiment et l'ombre car ils ont un meilleur contraste que les contours entre le sol et l'ombre) et l'autre de l'approximation des contours des régions sombres avec des segments (polygonalisation).

Les ombres sont extraites dans [Tison 2004] par une segmentation markovienne avec un grand nombre de classe. La classe présentant le niveau d'amplitude le plus faible constitue les ombres.

3.5.3.2. Extraction des échos et points brillants

Les algorithmes d'extraction peuvent être séparés en deux catégories suivant l'utilisation ou non des statistiques propres aux images RSO.

Approches adaptées à l'imagerie RSO

Dans [Tupin 1998], un détecteur de contour adapté à loi Rayleigh généralisée, est utilisé dans huit directions. La meilleure réponse du filtre est conservée et donne ainsi la direction du contour. Un post-traitement pour supprimer les fausses alarmes est effectué par seuillage sur la réponse du filtre et suppression des petites surfaces de moins de dix pixels. La valeur du seuil est obtenue par calcul d'après la valeur de la probabilité de fausse alarme désirée. Un masque binaire est alors obtenu. Cette détection est également utilisée dans [Simonetto 2002]. L'auteur effectue la fusion de ce masque avec un autre masque issu d'une classification reposant sur les chaînes de Markov cachées (cf. annexe C). Pour extraire les échos forts à partir des détections, l'auteur effectue, après mise en correspondance des primitives, une extraction simultanée sur les deux images par AROCS (Algorithme de reconnaissance des formes par optimisation d'un critère surfacique) en considérant comme primitives les demi-couronnes pour les cuves et les L formés par les échos des signatures des bâtiments. L'extraction simultanée des primitives sur les deux images est plus robuste.

Dans [Quartulli 2004], les auteurs modélisent les lois statistiques des images RADAR par l'inverse d'une gaussienne généralisée (loi à trois paramètres). Dans [Frery 1999a] et [Frery 1999b], les auteurs montrent que cette loi permet d'estimer les lois aussi bien pour les zones homogènes (clutter) que les réflexions fortes. Cette probabilité est exprimée dans un champ de Gibbs sous forme d'un potentiel. Associée à des informations plus haut niveau, l'extraction se fait en utilisant le critère du maximum *a posteriori*.

Autres approches

Dans [Michaelsen 2002] et [Stilla 2003b], les auteurs s'intéressent à l'extraction des points brillants sur les images en amplitude grâce à un détecteur emprunté à l'infrarouge ([Kohnle 1993]). Il s'agit d'un filtre composé du pixel central et de 3 « couronnes » de rayon 1, 2 et 3 autour de ce pixel. Sur chaque couronne, la valeur maximale du pixel est conservée. La sortie du filtre est obtenue par la différence d'amplitude du pixel central et du minimum des 3 maxima issus des couronnes. Le groupement perceptuel est ensuite utilisé pour associer les points brillants et obtenir ainsi des informations sur la forme des bâtiments à extraire en formant des lignes de points brillants selon les directions privilégiées. Les auteurs comptent sur la présence (régulière) de systèmes métalliques (aération ou autre) sur les toits des bâtiments. Deux seuils sont utilisés pour former les lignes de points brillants. Le premier (seuil bas) sert à déterminer si le point brillant va prendre part au groupement. Le second (seuil haut) sert à déterminer si le point brillant va servir d'initiateur pour le groupement

Dans [Petit 2004], la classification floue, effectuée sur l'image d'amplitude, permet d'obtenir deux classes composées des points brillants et très brillants. Cette deuxième classe consiste généralement en les coins réflecteurs.

Dans [Thiele 2006] l'extraction des échos est effectuée sur les deux images prises avec des trajectoires orthogonales. Les images d'amplitude géoréférencées sont fusionnées par un filtre maximum. Ensuite les structures linéaires sont extraites sur l'image fusionnée en tenant compte de leur épaisseur [Steger 1996]. Les lignes ayant une hauteur interférométrique similaire aux zones de sol environnantes sont considérées comme échos doubles.

Dans [Soergel 2003a], les échos forts sont obtenus des images en intensité par un seuillage puis un amincissement des contours allongés, connectés et brillants. La détection évolue dans [Soergel 2006] où les structures linéaires sont détectées sur les deux images, acquises avec des trajectoires orthogonales, par l'algorithme présenté dans [Foerstner 1994]. Cet algorithme calcule le gradient $G=[g_x, g_y]$ selon les deux axes de l'image. Les valeurs propres du produit GG^T sont ensuite calculées. Si une des deux valeurs propres a une valeur élevée, un contour est détecté. Un grand nombre de lignes est obtenu. Les espaces de quelques pixels entre les lignes colinéaires sont comblés.

3.5.4. Extraction de hauteurs

Selon le type de primitive et le type/nombre d'images de la même scène, les hauteurs peuvent être estimées par une ou plusieurs techniques, à savoir la monoscopie, la radargrammétrie et l'interférométrie. Dans ce paragraphe, sont détaillées les algorithmes de mises en œuvre de ces techniques.

3.5.4.1. Extraction de hauteurs à partir des ombres

L'extraction de la hauteur à partir des ombres de bâtiments en zone urbaine est délicate en imagerie RSO à cause des distorsions géométriques (repliement) et de la phénoménologie (réflexions triples) qui dégradent la qualité de l'extraction et fausse la mesure monoscopique. Dans [Stilla 2003a], les auteurs montrent que, pour une zone relativement dense, seulement 28% des ombres sont considérées comme fiables. Sur les données ERS ($\theta=23^\circ$) acquises sur Paris, le sol (les rues) dans les quartiers denses n'est pas visible. De plus, lorsque le sol est visible, les zones à faible rétrodiffusion telles que les routes lisses, apparaissent au même niveau d'amplitude que les ombres et dégradent également la fiabilité de cette mesure. Néanmoins elles restent d'un intérêt majeur dans les problématiques de détection, où

l'existence d'une ombre signifie la présence d'un bâtiment, et de reconstruction (la mesure monoscopique de la hauteur à partir de l'ombre s'avère la plus précise ; cf. §2.3.1).

Dans [Kirsht 1998], l'auteur détermine par monoscopie la hauteur des bâtiments (supposés à toit plat). Sans évaluation quantitative, l'article ne permet pas de conclure sur la précision de l'algorithme employé. L'utilisation de quatre traces orthogonales permet d'avoir des ombres sur tous les murs de l'édifice ([Bolter 2001]). Après extraction des ombres sur les quatre images grâce au filtre circulaire présenté au §3.5.3.1, les hauteurs des façades sont estimées par mesure monoscopique. Le bâtiment est délimité par quatre segments correspondant aux quatre backedges. Dans l'approche choisie par l'auteur, les bâtiments sont supposés à empreinte rectangulaire. Le rectangle passant par les quatre segments extraits est utilisé comme empreinte au sol. L'auteur constate qu'utiliser les ombres aboutit à une localisation des contours du bâtiment plus précise que celle effectuée sur l'interférogramme par la combinaison du maximum des quatre vues (cf. § 3.4.3.2). En effet, le repliement des façades ne permet pas de localiser précisément le bâtiment sur l'interférogramme.

La monoscopie sur les ombres, ligne par ligne permet également de déterminer la géométrie du toit ([Bolter 2001]). Il s'agit de la technique « shape-from-shadow ». Néanmoins, les ombres, parfois peu contrastées, sont difficiles à extraire et la mesure monoscopique sur une seule ligne se révèle bruitée.

3.5.4.2. Extraction de hauteurs à partir des échos forts

Trois techniques de mesures peuvent être utilisées de façon complémentaires pour extraire une information sur la hauteur des bâtiments à partir des échos forts : la monoscopie, la radargrammétrie et l'interférométrie.

La mesure monoscopique permet d'obtenir une mesure relative de hauteur. La hauteur absolue du bâtiment est alors obtenue par utilisation du modèle numérique de terrain ou de la radargrammétrie. Si plusieurs images de la même scène sont disponibles, il est possible d'obtenir plusieurs mesures monoscopiques. Sur un même ensemble de primitive sur chaque image, pour des raisons de précision, la mesure de la parallaxe sur l'image d'incidence la plus faible est privilégiée ([Simonetto 2002]).

A partir des échos extraits par AROCS sur les deux images d'amplitude, la technique employée pour mesurer la parallaxe diffère selon le bâtiment à reconstruire. Dans le cas de cuves, la distance entre les deux points brillants au milieu des deux demi-couronnes permet d'obtenir la hauteur de la cuve. Pour les bâtiments, l'auteur utilise l'écho du bord du toit le plus éloigné du capteur et effectue une détection des croix formées par l'écho du bord du toit et l'écho double lorsqu'ils sont visibles. L'auteur obtient ensuite, par mesure des segments l_1 et l_2 , la longueur s (cf. figure 60).

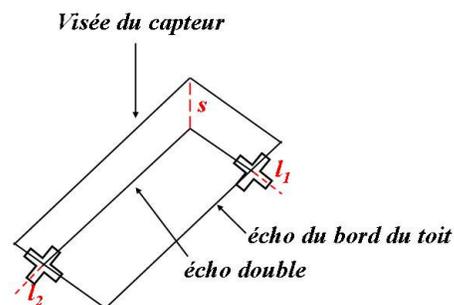


Figure 60 –Extraction de la mesure monoscopique sur les bâtiments [Simonetto 2002]

La deuxième approche consiste à utiliser la radargrammétrie lorsque deux images de la même scène sont disponibles. Dans [Simonetto 2002], les deux images sont issues d'acquisitions parallèles. Les demi couronnes des cuves et les ' L ' des bâtiments sont mis en correspondance, par corrélation (cf. §3.4.1.2), en appariant les points d'accès aux primitives (points brillants pour les cuves, croix correspondant à l'intersection des échos doubles et du bord du toit pour les bâtiments).

Dans [Soergel 2006] les deux acquisitions ont des trajectoires orthogonales. Le repliement sur les deux images est également selon des directions orthogonales. L'auteur n'a pas de données interférométriques et ne connaît pas la hauteur des L extrait. C'est pourquoi il les projette en géométrie terrain pour une fourchette de hauteur. Lorsque deux L projetés, appartenant à des images distinctes, se recouvrent, cela veut dire qu'ils correspondent au même coin du bâtiment dans l'espace 3D.

La dernière approche est interférométrique. La hauteur interférométrique des échos forts simples est utilisée pour estimer la hauteur du bâtiment. Néanmoins, comme nous l'avons montré au §2.5, le repliement doit être pris en compte. Dans [Thiele 2006], les échos doubles, à la hauteur du sol, permettent de délimiter l'empreinte au sol du bâtiment, supposé rectangulaire.

3.5.5. Reconstruction

Les algorithmes de reconstruction sont séparés en deux approches (ascendantes / descendantes).

3.5.5.1. Approche ascendante

L'approche ascendante consiste à reconstruire le bâtiment à partir d'informations bas niveau sur les images. Ces informations peuvent être sous forme de primitives ou de surfaces (les zones élevées ; cf. §3.4.3).

A partir de primitives

L'extraction des primitives sur les images RSO est une étape délicate et des erreurs d'extraction peuvent survenir. Les extractions peuvent également être incomplètes.

Le problème de la reconstruction devient alors un problème mal-posé : à partir d'un même ensemble de primitives, plusieurs bâtiments pourraient être reconstruits. Pour palier cette difficulté, la gestion d'hypothèse est employée. Elle se divise en deux étapes. La première consiste à générer tous les bâtiments possibles en fonction des primitives extraites et des modèles choisis. La seconde étape, la validation, sélectionne l'hypothèse optimisant un critère à définir.

Dans [Thiele 2006], l'auteur extrait les structures brillantes en L à partir des deux images acquises à des trajectoires orthogonales. Une sélection des L valides est effectuée en tenant compte des trajectoires des avions. En effet, les bâtiments sont supposés rectangulaires et l'angle entre les deux segments du L projetés en géométrie terrain, doit être un angle droit. A partir des L extraits, des hypothèses de bâtiments rectangulaires sont générées. Les hypothèses sont sélectionnées en deux étapes. La première sélection consiste à ne garder que les empreintes de bâtiment dont la hauteur moyenne est supérieure à celle du sol de quelques mètres et dont l'écart-type altimétrique doit être inférieur à un seuil non précisé. La seconde étape tient compte de la visibilité des L qui dépend des directions d'illumination. Si, à l'issue de cette sélection, des polygones ont des parties communes, celui qui est le plus probable est

conservé. Le critère utilisé dépend des hauteurs moyennes et des écart-types altimétriques sur chaque polygone mais il n'est pas détaillé. Seul le modèle de bâtiment à toit plat est envisagé. Il n'y a donc pas de gestion d'hypothèses effectuée sur le type de toit des bâtiments.

Dans [Soergel 2006], l'extraction des L sur le couple d'images radargrammétriques aboutit à de nombreuses fausses alarmes. La sélection d'hypothèse est effectuée lors de la projection des L en géométrie terrain. Toutes les associations de ces structures sur les deux images sont testées et seules celles formant des coins 3D sont conservées. Une mesure de hauteur moyenne du bâtiment est alors obtenue. Connaissant la hauteur moyenne du bâtiment, les L appartenant à un même bâtiment sont projetés en géométrie terrain. Les axes de symétrie du bâtiment sont alors déterminés.

A partir de zones élevées

Lorsque l'étape de reconstruction est envisagée sur les zones élevées, l'utilisation des modèles géométriques permet de réduire le bruit sur la mesure interférométrique. Ainsi, pour affiner les résultats fournis par la segmentation de l'interférogramme par l'algorithme des « scan lines », dans [Gamba 1999, 2000a, 2000b], les bâtiments sont supposés à toits plats. Le toit des bâtiments est estimé par le plan horizontal qui minimise l'erreur quadratique moyenne avec l'interférogramme. L'erreur nominale du capteur est de 2.5m. L'erreur moyenne sur l'extraction est de 2.2m avec un écart-type de 4.9m. Les zones élevées proches en hauteur (moins de 7.5m pour des bâtiments entre 40m et 110m) sont également fusionnées.

A partir de primitives et de zones élevées

Lors de la reconstruction, la première étape consiste généralement à déterminer l'empreinte au sol des bâtiments. La seconde détermine la forme du toit.

Dans [Soergel 2003a], l'auteur travaille sur des données interférométriques. Les contours issus des ombres et des échos forts sont associés en polygones en suivant les règles suivantes :

- au moins un contour de l'objet doit appartenir aux échos forts.
- un contour doit provenir du set de contour entre le toit et l'ombre.

Toutes les empreintes possibles sont générées et l'intersection de ces polygones et des régions élevées (extraites sur l'interférogramme), permet d'écarter les fausses alarmes.

Lorsque les empreintes sont déterminées, si des bâtiments avec différents modèles de toits sont utilisés, ils sont tous générés. La validation est effectuée par la comparaison des simulations des hypothèses de bâtiments et des données interférométriques. Pour chaque hypothèse, les plans caractéristiques des modèles sont calculés. Par exemple, pour un bâtiment avec un toit à double pentes, deux plans sont utilisés. Cette étape peut être automatique ([Bolter 2001], [Soergel 2003a]) ou semi-automatique ([Baalz 2003]) : l'utilisateur sélectionne le modèle et son orientation. Les résultats de la simulation sont alors comparés aux primitives extraites et peuvent guider une nouvelle extraction pour améliorer les résultats. Le critère de sélection porte sur l'hypothèse présentant [Soergel 2003a]:

- le meilleur recouvrement et parallélisme par rapport aux primitives extraites.
- le meilleur rapport zone élevée sur aire du polygone.

Dans [Petit 2004], l'intersection de la classe floue « coins réflecteurs » et du masque correspondant au sursol permet de déterminer la localisation des constructions urbaines. Une croissance de région sur l'interférogramme permet alors d'obtenir un masque de bâtiment qui

est ensuite affiné par l'utilisation des modèles de bâtiments rectangulaires. Un bâtiment peut être décrit par la réunion de plusieurs rectangles. Les hauteurs des toits des bâtiments ne sont pas moyennés mais débruités par un filtre médian.

3.5.5.2. Méthode descendante

La méthode descendante consiste à considérer la reconstruction d'une façon diamétralement opposée. Chaque modèle est caractérisé par des paramètres. Dans [Quartulli, 2001], [Quartulli, 2003a], [Quartulli, 2004], le bâtiment est décomposé en facettes, chacune caractérisée par leur centroïde, dimensions et réflectivité RADAR. Les auteurs utilisent une approche probabiliste basée sur le critère du maximum *a posteriori* pour déterminer les paramètres de chaque facette.

Pour cela elles sont modélisées de façon probabiliste par un ensemble de pixels dans l'image d'amplitude. Les auteurs écrivent alors la probabilité *a posteriori* d'obtenir la scène en fonction des bâtiments présents, sous forme d'un champ de Gibbs. Le potentiel est décomposé en :

- un terme d'*a-priori*, dépendant des primitives objets, et permettant de tenir compte des recouvrements entre les objets.
- un terme de ressemblance, dépendant des pixels et permettant d'extraire les échos forts.

Il n'y a pas d'expression analytique pour l'expression de la probabilité *a posteriori* et l'optimisation se fait par les méthodes de Monte Carlo (échantillonnage de Gibbs sur un seul paramètre à la fois) avec un recuit simulé [Ortner 2002]. Lors de l'optimisation, le passage d'un type de bâtiment à un autre est autorisé.

Sur les images tests, tous les bâtiments sont détectés mais ne sont pas tous bien reconstruits. Deux bâtiments proches et colinéaires ne sont pas séparés. La largeur des bâtiments selon l'axe des distances est parfois approximative.

3.5.5.3. Algorithmes itératifs

La reconstruction directe de bâtiments complexes est délicate car les primitives RADAR sont difficiles à extraire et à associer. Ainsi de nombreux auteurs adoptent une approche itérative utilisant un simulateur géométrique. Les algorithmes itératifs se révèlent être une combinaison d'approche ascendante et descendante. La simulation géométrique consiste à projeter en géométrie RADAR les hypothèses obtenues et de descendre vers le bas niveau. Cette descente permet alors de valider ou de rejeter certaines hypothèses. L'utilisation d'une nouvelle étape ascendante, avec cette fois un *a-priori* fort sur le bâtiment à reconstruire, permet d'améliorer la reconstruction à l'itération suivante.

L'itérativité peut consister en un recuit simulé [Quartulli 2004] et trouver ainsi la solution globale qui maximise la probabilité *a posteriori*. Elle peut également être utilisée pour minimiser une erreur quadratique moyenne [Soergel 2003a]. Ainsi, lors de la première itération, ne sont modélisés que des bâtiments rectangulaires. Les dépendances sont reconstruites lors d'itérations ultérieures. L'orientation des bâtiments est également corrigée. L'algorithme s'arrête lorsque l'erreur quadratique moyenne entre les données simulées et les données réelles est faible. En utilisant deux vues de traces orthogonales, l'erreur de localisation obtenue est de 3m. Celle sur l'élévation est inférieure à 1.5m.

L'utilisation de l'approche itérative permet également d'affiner la reconstruction par détection des occultations et du repliement [Bolter 2001]. Après fusion des informations planimétriques, sur les images en amplitude, et altimétriques, grâce à la combinaison du

« maximum des quatre vues », une simulation est effectuée. Comparées aux images réelles, l'auteur peut alors nettoyer les ombres. De la même façon, sur les images interférométriques, l'auteur détecte les zones de repliement. Elles sont ensuite masquées afin de calculer, avec une meilleure précision altimétrique, la hauteur du bâtiment.

3.5.6. Conclusion

Le principe de la reconstruction, utilisant des modèles de bâtiments a été présenté. Les algorithmes de reconstruction à partir de primitives ou des zones élevées ont été divisés selon deux approches (ascendante et descendante). A cause des erreurs d'extraction et de la non-exhaustivité des extractions, la gestion d'hypothèses est envisagée pour générer un ensemble d'hypothèses de bâtiments. La sélection est ensuite effectuée sur des critères surfaciques (type de toit ou hauteur minimale) ou géométriques (appariement de primitives dans le repère monde, orientation des primitives). La reconstruction de bâtiments complexes (avec des dépendances, formes de toits diverses) s'avère une étape difficile et est effectuée grâce à des algorithmes itératifs.

3.6. Conclusion

L'état de l'art sur la représentation en trois dimensions de bâtiments en imagerie RSO a été présenté. Les différents algorithmes de traitement des images et d'extraction de MNS, MNT, zones élevées ou primitives (ombres, échos) sur les images RSO ont été détaillés. Indépendamment de la nature des images traitées, ils se divisent en trois catégories qui traduisent des approches différentes.

La première se rapproche du signal et regroupe des algorithmes tenant compte des particularités statistiques de l'imagerie RADAR :

- classification par chaînes ou champs de Markov adaptées aux lois de Rayleigh ou Fisher
- détecteur de lignes adaptées à la loi de Rayleigh
- extraction d'échos fort utilisant la loi inverse de la gaussienne généralisée
- débruitage grâce à des filtres paramétriques

Elles ont l'avantage de modéliser théoriquement les primitives à extraire.

La seconde, basée sur des algorithmes provenant d'autres domaines (optique, infrarouge), essaye de ramener les images dans des conditions adéquates d'utilisation par un préfiltrage tel que le multivue (entre autres) :

- algorithme des scan-lines
- croissance de régions
- détecteurs de points brillants
- mise en correspondance de primitives sur un couple radargrammétrique par corrélation sur les moyennes ou variances
- extracteur de lignes par [Steger 1996] ou [Forstner 1994]

Ces méthodes permettent de réutiliser des algorithmes déjà développés. Par contre, les hypothèses d'utilisation ne sont plus toujours valides.

La dernière regroupe des algorithmes non paramétriques :

- filtre médian
- filtre rotatif circulaire pour l'extraction des ombres
- seuillage pour l'extraction des échos

- morphologie mathématique

Ces filtres présentent l'avantage de ne pas nécessiter de modélisations statistiques et sont applicables à tout type d'images.

Les méthodes d'extraction de la hauteur et de la localisation des bâtiments ont également été détaillées. Comme nous l'avons montré au §2.3, les mesures monoscopiques se révèlent précises d'un point de vue altimétrique et planimétrique. Ce résultat est confirmé pour les ombres dans les travaux de [Bolter 2001].

L'extraction des hauteurs pour un couple d'images radargrammétriques impose le géoréférencement des images. La mise en correspondance des primitives de l'image peut être effectuée sur les images en considérant des critères portant sur les amplitudes ou simplement sur la localisation en géométrie terrain. L'utilisation des données interférométriques est un point fort de l'imagerie RSO pour l'obtention du 3D car elle permet d'avoir facilement une estimation de la hauteur des primitives ou de zones élevées. En revanche cette estimation est peu précise et les limitations de l'interférométrie pour la reconstruction de bâtiments dues au repliement ont été mises en évidence par Bolter.

Enfin, le principe de la reconstruction basé sur l'utilisation de modèles géométriques a été présenté. Les modèles sont construits à partir d'hypothèses effectuées sur les bâtiments à reconstruire. Deux approches sont distinguées. L'approche ascendante utilise les informations extraites sur les images pour aboutir à un bâtiment conforme aux modèles. Les algorithmes basés sur cette approche utilisent les primitives extraites et/ou les zones élevées pour obtenir les hypothèses de bâtiments. Les bâtiments reconstruits par certains des algorithmes proposés ont des signatures RSO très compliquées [Soergel 2006]. Ce type d'approche semble donc prometteur. L'approche descendante consiste à projeter un modèle en géométrie RADAR et à estimer les paramètres de ces modèles d'après la signature obtenue. On peut remarquer que les approches ascendantes sont nettement plus nombreuses car elles sont plus intuitives.

L'utilisation d'un simulateur au sein d'algorithmes itératifs permet de combiner les approches ascendantes et descendantes et d'améliorer la précision sur la reconstruction.

**Partie 2 : Proposition d'une méthode de
reconstruction de bâtiments sur images RSO
interférométrique haute résolution**

Lors de l'étude bibliographique, les approches utilisant la gestion d'hypothèses sur données interférométriques pour la reconstruction de bâtiments ont été présentées (cf. §3.5.5). La plupart des méthodes existantes utilisent les échos forts et/ou les ombres pour délimiter les bâtiments sous forme de parallélogramme [Soergel 2003a], [Quartulli 2004], [Thiele 2006], [Bolter 2001]. Le type de modèle et les hauteurs du toit peuvent ensuite déterminés à partir de ces parallélogrammes en utilisant l'interférogramme [Soergel 2003a], [Thiele 2006]. Les ombres sont rarement utilisées pour déterminer la hauteur des bâtiments car elles sont considérées comme peu fiables [Stilla 2003]. Néanmoins l'algorithme de [Bolter 2001] basé sur la reconstruction par les ombres fournit des résultats intéressants grâce à la bonne précision de la mesure monoscopique.

De plus un algorithme se basant principalement sur l'interférogramme pour le calcul des hauteurs [Gamba 2000], [Thiele 2006] serait peu fiable sur les images RAMSES en bande X car les effets du repliement sur l'interférogramme ont été mis en évidence au §2.5. Les données interférométriques sont un atout majeur pour la reconstruction car elles fournissent une information dense de hauteur mais elles doivent être utilisées avec précaution après étiquetage des zones de repliement.

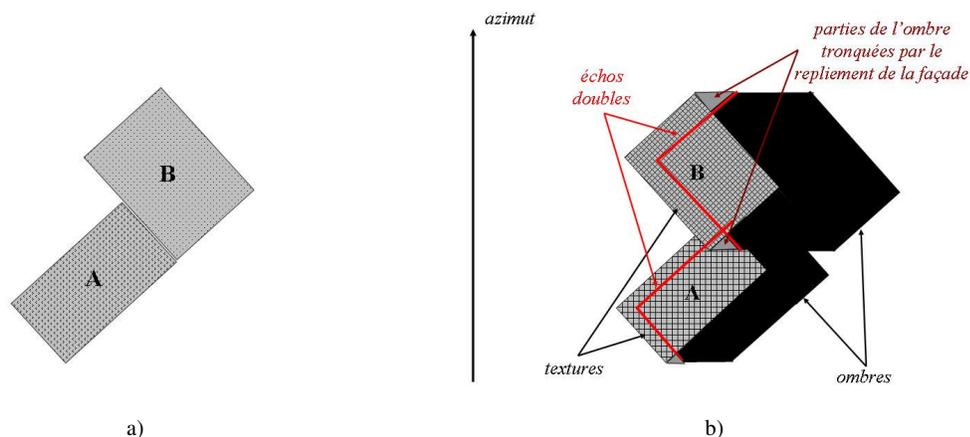
La majorité de ces algorithmes utilisent plusieurs acquisitions de la même zone. Si une seule acquisition est disponible, les résultats sont moins bons [Soergel 2003a] ou les algorithmes sont tout simplement inutilisables [Thiele 2006].

Dans cette partie, une méthode de reconstruction de bâtiments sur image RSO haute résolution en bande X est proposée. La reconstruction est effectuée à partir d'une seule acquisition interférométrique du capteur RAMSES sur une zone industrielle. Moins d'informations sont disponibles par rapport aux algorithmes qui utilisent plusieurs acquisitions de la même zone. Nous nous situons donc dans un cadre d'application très difficile. Dans notre travail, nous choisissons de nous restreindre à la reconstruction de bâtiments et non à la détection. C'est pourquoi les zones utilisées consistent en de petites images avec un bâtiment isolé.

Les algorithmes « image-based » utilisant des modèles géométriques de bâtiments ont montré leur capacité à réduire le bruit sur les mesures interférométriques. Cette étude s'inscrit dans cette catégorie. Les hypothèses choisies sur les bâtiments sont des empreintes à angles droits et des toits plats ou plats par morceaux (en escalier). Elles suffisent à décrire un large spectre de bâtiments. Nous nous proposons d'utiliser, pour la reconstruction :

- les échos forts, avec prise en compte du repliement pour le calcul des hauteurs
- les textures
- les ombres, avec détermination de leur fiabilité
- les façades avant (vues par le capteur) et arrière (non vues par le capteur) grâce à un algorithme de détection
- l'interférogramme, avec estimation de la hauteur des toits en tenant compte du repliement

Sur la figure 61a, un bâtiment en L est représenté schématiquement. Ce bâtiment est divisée en deux parties, notées A et B. La partie B est choisie plus haute que la partie A. Un exemple de signature, lorsque les échantillonnages en distance et en azimut sont les mêmes, est proposé en figure 61b. Dans ce cas les angles droits des bâtiments sont conservés. Sur les deux toits apparaissent des textures. Les échos doubles, caractérisant l’empreinte au sol du bâtiment, sont représentés en rouge. Les ombres, partiellement tronquées par le repliement des façades supposées rétrodiffusantes, sont en noir.



Les échos doubles et les textures des toits suivent les deux orientations du bâtiment. Les éventuelles textures associées à la rétrodiffusion sur les façades ne sont pas prises en compte. Les ombres suivent également ces deux directions et celle associée à l’axe des distances (lorsque le doppler centroïde est nul).

La première étape de cette chaîne consiste en l’extraction des signatures des bâtiments. Néanmoins, nous avons vu que les primitives (échos, ombres et textures) associées aux bâtiments sont fortement bruitées. Les images sont tout d’abord filtrées pour réduire l’aspect granuleux du chatoiement. Une extraction automatique des signatures semble difficile car les primitives présentent des amplitudes variables. C’est pourquoi les étapes de détection et d’extraction sont guidées par l’utilisateur. Pour améliorer l’extraction, l’algorithme proposé utilise les hypothèses choisies sur les modèles de bâtiments à reconstruire en considérant les primitives à extraire comme des segments ou des surfaces avec des contours orientés suivant les deux directions du bâtiment projeté en géométrie RADAR et l’axe des distances.

Les signatures des bâtiments sur les images RSO sont complexes. Souvent, une analyse au cas par cas est nécessaire pour les comprendre dans sa globalité. Il paraît alors peu réaliste d’envisager une reconstruction directe aboutissant à un unique modèle. C’est pourquoi, une approche basée sur la génération / sélection d’hypothèses est choisie. Contrairement aux approches proposées par [Soergel 2006] ou [Thiele 2006], les jeux de primitives échos, ombres et textures utilisées lors de la reconstruction sont considérés comme exacts mais incomplet. Nous voulons voir jusqu’où peut aller la reconstruction, en limitant le nombre d’hypothèses générées, par une approche « imaged-based » à partir de ces primitives.

4. Extraction de signatures de bâtiments en imagerie RSO haute résolution

4.1. Introduction

L'extraction de signatures de bâtiments en imagerie RSO haute résolution est un point dur de la chaîne de reconstruction. En effet, le chapitre 2 montre la variabilité des signatures. Les textures sur les toits des bâtiments ne sont pas toujours présentes. Les ombres peuvent être bruitées par des échos triples ou tronquées à cause du repliement. Les échos simples ou doubles n'apparaissent pas de façon systématique. De plus, si ces échos sont présents, ils ne possèdent pas le même niveau d'amplitude d'un bâtiment à l'autre. Une approche complètement automatique apparaît donc difficilement fiable.

Dans ce chapitre sont détaillées les étapes principales de l'extraction des signatures : le filtrage, la détection et la caractérisation. La première étape pour l'extraction des primitives est le filtrage permettant de réduire le chatoiement. Les filtres usuels sont paramétriques : il tiennent compte de la nature statistique particulière des images RSO en amplitude. Le filtre non-paramétrique mean shift, habituellement utilisé en imagerie optique, est introduit en imagerie RSO et comparé aux filtres usuels. Ce filtre est également étendu au filtrage des ombres en tenant compte de l'information contenue dans l'image de cohérence. A l'issue de la détection des prémices de primitives (ombres, échos et textures) plus ou moins bruités sont obtenus. Néanmoins, ils sont difficilement utilisables directement pour la reconstruction. L'étape d'extraction permet, grâce à des hypothèses sur la forme des toits et sur l'empreinte au sol des bâtiments à reconstruire, d'obtenir un jeu de primitives adaptées à la reconstruction.

4.2. Introduction du filtre mean shift en imagerie RSO

4.2.1. Nécessité d'un filtrage des images RSO en amplitude

La statistique de Rayleigh-Nakagami des images RSO en amplitude est utilisée pour modéliser le chatoiement (cf. annexe B.2.1). Une zone homogène, comme le toit du bâtiment présenté sur la figure 24a, p43, apparaît granuleuse. Le chatoiement est d'autant plus important que le niveau de rétrodiffusion ou réflectivité est élevé. Ces fluctuations d'amplitude gênent les algorithmes de détection ou d'extraction traditionnels. Par exemple, le toit de ce bâtiment n'a pas pu être extrait par un algorithme de croissance de région. Une étape de filtrage est donc nécessaire.

Les algorithmes de filtrage peuvent utiliser les statistiques sous-jacentes au sein des images. L'estimation des paramètres des lois statistiques peut se décomposer en deux catégories. La première consiste à utiliser des connaissances *a priori*, telles que le type de loi présente sur les images, et à estimer leurs paramètres par des méthodes telles que le maximum de vraisemblance, la méthode des moments ou le maximum *a posteriori*. On parle alors d'estimation paramétrique de lois statistiques. La deuxième catégorie, dite non-paramétrique,

permet de faire l'estimation des lois sans avoir à passer par l'estimation de ses paramètres. Les méthodes à noyaux, utilisées par le mean shift, s'inscrivent dans cette deuxième catégorie. Dans ce paragraphe, nous étudions l'application du filtre mean shift en imagerie RSO.

4.2.2. Filtres paramétriques usuels en imagerie RSO

Pour estimer la réflectivité locale, les filtres RSO paramétriques tentent de minimiser un critère d'erreur quadratique moyenne (filtres de Lee, Kuan, Frost) ou de maximiser une probabilité *a posteriori* (filtres GammaMAP, FisherMAP). Pour tous ces filtres la taille de la fenêtre est déterminante dans le résultat du filtrage. Trop peu de pixels et les paramètres des statistiques seront mal estimés. Trop de pixels et les estimations seront fausses car effectuées sur plusieurs régions de statistiques différentes. Ce paragraphe est basé sur [Maître 2001].

4.2.2.1. Filtres basés sur la minimisation de l'erreur quadratique moyenne (MEQM)

Lee et Kuan : ces deux filtres utilisent le modèle multiplicatif : $i = R.S$, où i désigne l'intensité des pixels de l'image, S le chatoiement et R la réflectivité à estimer (cf. annexe B.2.1.3). Ces filtres sont adaptatifs linéaires, c'est à dire qu'ils estiment la réflectivité par :

$$\hat{R} = E(i) + k(i - E(i))$$

où $E(\cdot)$ désigne l'espérance mathématiques.

Il faut alors déterminer k pour ces deux filtres sachant que le filtre de Lee utilise l'hypothèse :

$$(R - E(R))(i - E(i)) \approx 0.$$

Le coefficient de variation γ est défini par $\gamma^2 = \frac{m_2 - m_1^2}{m_1^2}$ où m_1 et m_2 sont les moments d'ordre 1 et 2. Pour ces deux filtres k s'exprime en fonction des coefficients de variation de l'intensité (γ_I) et du speckle normalisé ($\gamma_S = \frac{1}{\sqrt{L}}$) où L est le nombre de vues.

Pour Lee :

$$k = 1 - \frac{\gamma_S^2}{\gamma_I^2}$$

Pour Kuan :

$$k = \frac{1 - \frac{\gamma_S^2}{\gamma_I^2}}{1 - \gamma_S^2}$$

Ces deux filtres ont un comportement très proche : lorsque la zone est homogène, $\gamma_S^2 \approx \gamma_I^2 \Rightarrow k \approx 0$ et la valeur moyenne de la zone est choisi comme meilleur estimateur de la réflectivité. Lorsque la zone est fortement hétérogène, $\gamma_S^2 \ll \gamma_I^2 \Rightarrow k \approx 1$ et la valeur du pixel est conservée.

Pour les lois Gamma généralisée et Rayleigh généralisée, le coefficient de variation vaut respectivement $\frac{1}{\sqrt{L}}$ et $\sqrt{\frac{Poch(L+0.5,0.5)}{Poch(L,0.5)} - 1}$ [Nicolas 2003] où *Poch* désigne la fonction de Pochhammer.

Frost : ce filtre reprend les hypothèses des filtres de Lee et Kuan mais tient compte également de la corrélation de la scène. Plus le pixel est loin du centre de la fenêtre et moins

sa valeur aura de poids dans l'estimation de la réflectivité. La réponse impulsionnelle de ce filtre est :

$$m(d) = K\alpha \exp(-\alpha d) \text{ avec } \alpha^2 = K'\gamma_i^2 \text{ et } d \text{ la distance du pixel au centre de la fenêtre.}$$

- K est une constante de normalisation (somme des poids du voisinage)
- K' permet de régler l'importance du filtrage. Une grande valeur filtrera faiblement le voisinage que la zone soit hétérogène ou non. Une faible valeur filtrera plus la zone.

Dans [Lopès 1990], ces filtres sont améliorés en utilisant deux seuils sur le coefficient de variation de la fenêtre de filtrage. Si le coefficient de variation calculé est

- inférieur au seuil bas (C_{min}), la moyenne estime la réflectivité du pixel.
- supérieur au seuil haut (C_{max}), la valeur du pixel est gardée
- entre les deux, le filtrage s'opère.

Les résultats donnés par les filtres sont meilleurs mais nécessite l'utilisation de deux paramètres supplémentaires.

Wu et Maitre : pour éviter le choix de la fenêtre de calcul, ce filtre effectue une première étape de croissance de région à partir d'une zone carrée. Le critère d'homogénéité de cette croissance est le coefficient de variation. Le filtrage est ensuite effectué sur cette zone par un filtre de Kuan.

4.2.2.2. Filtres basés sur le critère du maximum à posteriori (MAP)

La deuxième catégorie de filtres paramétriques est basée sur la statistique bayésienne. Ces filtres tentent de maximiser la probabilité d'obtenir la réflectivité R connaissant l'intensité i :

$$p(R|i) = \frac{p(i|R)p(R)}{p(i)} \sim p(i|R)p(R)$$

Maximiser $p(R|i)$ revient à maximiser $p(i|R)p(R)$. Ce maximum est atteint lorsque la dérivée logarithmique de $p(i|R)p(R)$ est nulle.

$$\frac{1}{p(i|R)} \frac{\partial p(i|R)}{\partial R} + \frac{1}{p(R)} \frac{\partial p(R)}{\partial R} = 0 \quad (66)$$

Dans le cas d'un chatoiement pleinement développé la probabilité d'avoir l'intensité i connaissant la réflectivité R s'exprime par [Nicolas 2004]:

$$p(i|R) = \left(\frac{L}{R}\right)^L \frac{i^{L-1}}{\Gamma(L)} \exp\left(-\frac{Li}{R}\right) \quad (67)$$

Selon le filtre, la loi statistique utilisée pour caractériser la texture de la réflectivité diffère.

GammaMAP : la loi Gamma généralisée est ici utilisée comme modèle statistique pour la texture sous-jacente.

$$p(R) = \frac{1}{\Gamma(M)} \frac{M}{\mu} \left(\frac{MR}{\mu}\right)^{M-1} \exp\left(-\frac{MR}{\mu}\right) \quad (68)$$

La loi résultante du produit $p(i|R)p(R)$ est la loi K (cf. équation (107), p210). La résolution de l'équation (66) en utilisant (67) et (68) donne :

$$\hat{R} = \frac{1}{2M} \left(\mu(M-L-1) + \sqrt{\mu^2(M-L-1) + 4ML\mu i} \right)$$

FisherMAP : la loi Gamma inverse est proposée comme modèle statistique pour la texture.

$$p(R) = \frac{1}{\Gamma(M)} \frac{1}{M\mu} \left(\frac{M\mu}{R}\right)^{M+1} \exp\left(-\frac{M\mu}{R}\right) \quad (69)$$

Le produit $p(i|R)p(R)$ conduit à la loi de Fisher (cf. équation (108), p210). La résolution de l'équation (66) donne :

$$\hat{R} = \frac{Li + M\mu}{1 + L + M} \quad (70)$$

4.2.3. Le filtre Mean Shift

Le filtre mean shift [Fukanaga 1975] appartient à la catégorie des filtres non-paramétriques. Il utilise le lissage d'histogramme par méthode à noyaux (§4.2.3.1) et fonctionne de façon itérative pour déplacer la valeur de chaque échantillon vers le mode local de la statistique locale. En traitement d'image, le mean shift est utilisé en tenant compte du voisinage spatial [Commaniciu 2002]. Les diverses approches théoriques liées au mean shift sont résumées dans [Grenier 2005].

4.2.3.1. Lissage d'histogramme par méthode à noyaux

Les méthodes à noyaux sont utilisées dans le Mean Shift pour lisser des histogrammes locaux. Pour détailler le principe de ce lissage, une partie du chapitre 2 de [Silverman 1986] est présentée.

Soit χ une variable aléatoire de densité de probabilité (ddp) f^1 et Ω une réalisation de χ tel que $\text{card}(\Omega) = n$. $\exists(a,b) \in \mathfrak{R}^2$ tel que $\forall i \in [1..n]$, $\Omega(i) \in [a,b]$.

Après échantillonnage régulier du compact $[a,b]$ en m intervalles, l'ensemble des « bins » est défini comme étant :

$$\left\{ \left[a + \frac{k(b-a)}{m}, a + \frac{(k+1)(b-a)}{m} \right] \right\}_{k \in [0..m-1]}$$

On note alors h la largeur du bin. Celle-ci s'exprime par $h = \frac{b-a}{m}$.

Un estimateur usuel de cette ddp est l'**histogramme empirique** :

$$\hat{f}(x) = \frac{1}{nh} * (\text{nbre d'échantillons dans le même bin que } x)$$

Cette estimation se révèle très simple mais présente les inconvénients d'être discontinue et de changer fortement lorsque les centres des bins varient légèrement (cf. figure 62).

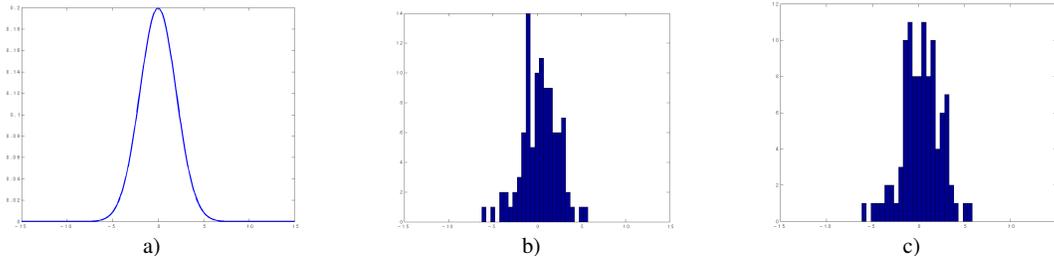


Figure 62 - a) ddp de la loi normale de moyenne nulle et d'écart-type 2, b) histogramme d'une réalisation de 100 échantillons de cette ddp avec une largeur de bin de 0.5, c) histogramme avec le centre des 'bins' décalés de 0.1 de cette même réalisation

¹ L'estimation par noyau ne requiert pas de propriété particulière sur la densité à estimer. Néanmoins, dans la suite de ce document elle sera continue et dérivable.

Pour palier ce dernier inconvénient un autre estimateur, dit, « naïf » peut être utilisé. Cet estimateur se ramène à la définition de $f(x)$:

$$f(x) = \lim_{h \rightarrow 0} \frac{1}{2h} P(x-h < X < x+h)$$

Il consiste à placer une boîte glissante de largeur $2h$ fixe et de hauteur $(2nh)^{-1}$ centrée en x . L'estimation de la ddp peut s'écrire alors :

$$\hat{f}(x) = \frac{1}{2nh} * (\text{nbre d'échantillons } x_i \text{ entre } x-h \text{ et } x+h)$$

En définissant le noyau uniforme w par l'expression : $w(x) = \begin{cases} 0.5 & \text{si } |x| < 1 \\ 0 & \text{si } |x| \geq 1 \end{cases}$, l'estimée de f s'écrit :

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} w\left(\frac{x-x_i}{h}\right)$$

Cet estimateur est connu sous le nom d'estimateur de Parzen et peut être utilisé avec d'autres noyaux (triangulaire etc...).

Un exemple d'estimation naïve d'une loi normale est donné figure 63. L'estimation est plus proche de la ddp théorique que dans le cas de l'histogramme mais les discontinuités, provenant de la non continuité de w en x_i-h et x_i+h , sont toujours présentes.

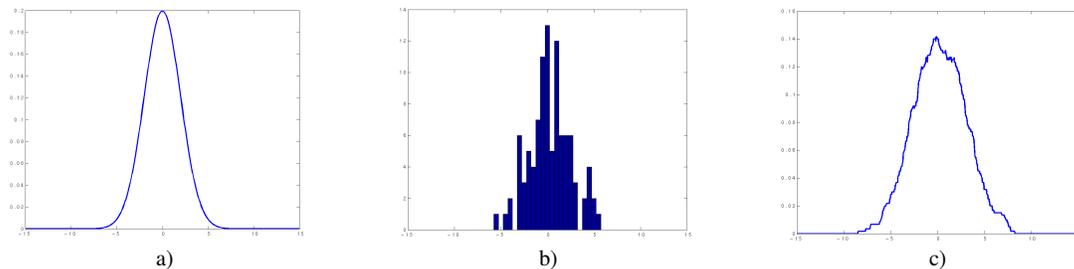


Figure 63 - a) ddp de la loi normale de moyenne nulle et d'écart-type 2, b) histogramme d'une réalisation de 100 échantillons de cette ddp avec une largeur de bin de 0.5, c) estimation naïve de la densité correspondante avec $h = 3$

Aussi pour estimer l'histogramme d'une ddp continue, l'utilisation d'un noyau continu semble plus adaptée. Ainsi, en remplaçant le noyau uniforme par un noyau normal par exemple, l'estimation de f héritera des propriétés de continuité et de dérivabilité de ce noyau (cf. figure 64). L'estimation de f peut s'écrire :

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x-x_i}{h}\right) \quad (71)$$

avec K_N une gaussienne de moyenne nulle et de variance unité

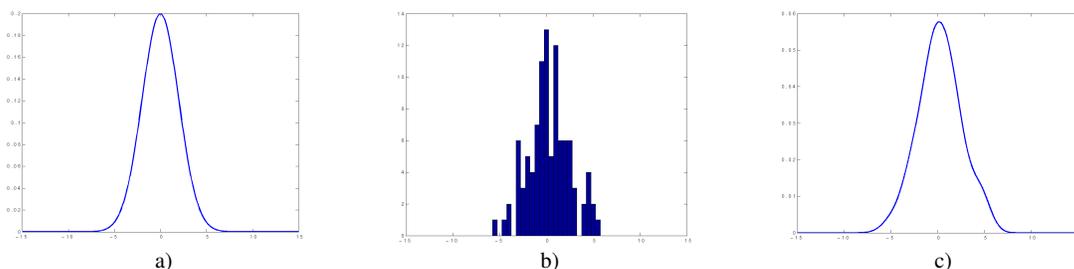


Figure 64 - a) ddp de la loi normale de moyenne nulle et d'écart-type 2, b) histogramme d'une réalisation de 100 échantillons de cette ddp avec une largeur de bin de 0.5, c) estimation par noyau normal de la densité correspondante avec $h = 3$

Dans le cas de noyaux continus, le paramètre h est appelé largeur de bande. Il est déterminant dans l'estimation par noyau. Son influence peut faire fortement varier l'allure de l'estimée de f . Une valeur de bande très faible pourra faire apparaître des modes qui n'existent pas tandis qu'une valeur trop grande pourra faire disparaître certains modes (cf. figure 65).

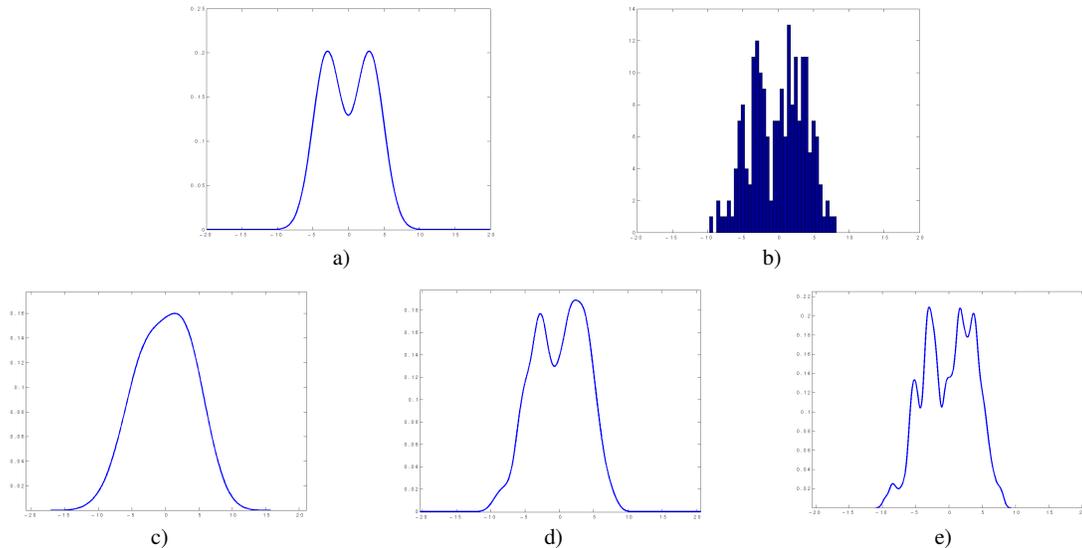


Figure 65 – a) ddp de deux lois normales de moyenne -3 et 3 et d'écart-type 2 , b) histogramme d'une réalisation de 100 échantillons de chaque loi avec une largeur de bin de 0.5 , estimation par noyau normal de la densité correspondante avec c) $h=8$, d) $h=3$ et e) $h=1.5$

4.2.3.2. Définition du Mean Shift

Etant donné un noyau K et n échantillons $\{x_1, x_2, \dots, x_n\}$, l'estimée \hat{f} de la ddp s'exprime :

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)$$

où $x \in [\min(\{x_1, x_2, \dots, x_n\}) - h, \max(\{x_1, x_2, \dots, x_n\}) + h]$.

En particulier, pour l'échantillon x_k , on a :

$$\hat{f}(x_k) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x_k - x_i}{h}\right)$$

Il est alors possible de calculer le « sample mean » en x correspondant à la moyenne des échantillons pondérés par le noyau K . Cette pondération dépend de la distance normalisée par la largeur de bande entre les échantillons x_i pour $i=1..n$ et x :

$$m_{h,K}(x) = \frac{\sum_{i=1}^n x_i K\left(\frac{x-x_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)}$$

Cette valeur n'a aucune raison d'être identique à la valeur x . La différence $m_{h,K}(x) - x$ est appelée le « mean shift » et est noté $ms_{h,K}$. On peut en particulier l'estimer pour l'échantillon x_k :

$$ms_{h,K}(x_k) = \frac{\sum_{i=1}^n x_i K\left(\frac{x_k - x_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_k - x_i}{h}\right)} - x_k \quad (72)$$

Il traduit l'écart entre l'échantillon x_k et la moyenne de tous les échantillons pondérée par le noyau K .

4.2.3.3. Principe de l'algorithme

L'algorithme mean shift a été développé par Fukanaga et Hostetler [Fukanaga 1975] pour faire une analyse par groupe de données similaires selon un critère à définir. Pour une image, par exemple, ce critère peut porter sur les niveaux de gris. Cet algorithme fonctionne de façon itérative et tend à faire converger chaque élément vers le mode local de la statistique estimée par la méthode des noyaux décrite au §4.2.3.1.

Le noyau utilisé est habituellement une ddp définie, symétrique et positive sur un intervalle donné [Wand 1994], [Comaniciu 2002]. De plus on suppose que le noyau K est symétrique. Il peut alors être défini par un « profil » noté k , tel que :

$$K(x) = c.k\left(\|x\|^2\right)$$

où c est une constante de normalisation.

Soient K un noyau satisfaisant ces conditions et $\{x_1^0, x_2^0, \dots, x_n^0\}$ le jeu de données à l'itération 0.

Au début de l'**itération m**, le jeu d'échantillon obtenu est noté $\{x_1^m, x_2^m, \dots, x_n^m\}$. Pour chaque x_k^m avec $k=1..n$, l'histogramme des échantillons est lissé par méthode à noyaux.

$$\hat{f}_{h,K}(x_k^m) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x_k^m - x_i^m}{h}\right) = \frac{1}{n} \sum_{i=1}^n \frac{c}{h} k\left(\left\|\frac{x_k^m - x_i^m}{h}\right\|^2\right)$$

Pour déterminer la position relative du mode de l'histogramme lissé par rapport à l'échantillon x_k^m il suffit de calculer la dérivée de l'histogramme lissé. Une pente positive indique la direction du mode le plus proche. En supposant que f soit continue et dérivable en chaque x_k^m avec $k=1..n$ et en utilisant la linéarité de la dérivée, l'estimée du gradient de f en x_k^m s'exprime par :

$$\hat{\nabla} f_{h,K}(x_k^m) \equiv \hat{\nabla} \hat{f}_{h,K}(x_k^m) = \frac{2}{n} \sum_{i=1}^n \frac{c}{h} (x_k^m - x_i^m) k'\left(\left\|\frac{x_k^m - x_i^m}{h}\right\|^2\right)$$

où k' est la dérivée de k par rapport à x .

$$\hat{\nabla}f_{h,K}(x_k^m) = -\frac{2c}{nh} \left[\sum_{i=1}^n k' \left(\left\| \frac{x_k^m - x_i^m}{h} \right\|^2 \right) \right] \left[\frac{\sum_{i=1}^n x_i^m k' \left(\left\| \frac{x_k^m - x_i^m}{h} \right\|^2 \right)}{\sum_{i=1}^n k' \left(\left\| \frac{x_k^m - x_i^m}{h} \right\|^2 \right)} - x_k^m \right] \quad (73)$$

En posant $k_F = k'$ et $K_F(x) = c_F k_F(\|x\|^2)$, l'expression du terme mean shift de l'équation (72) est retrouvée dans la seconde partie de l'équation (73), à savoir :

$$ms_{h,K_F}(x_k^m) = \frac{\sum_{i=1}^n x_i^m K_F \left(\frac{x_k^m - x_i^m}{h} \right)}{\sum_{i=1}^n K_F \left(\frac{x_k^m - x_i^m}{h} \right)} - x_k^m$$

où K_F est appelé kernel fantôme et est implicitement déterminé par le choix du noyau K .

Pour simplifier l'écriture, dans la suite de ce document, la notation suivante est adoptée :

$$ms_{h,K_F}(x) = MS_{h,K}(x)$$

En posant $g_{h,K}(x) = \left[\sum_{i=1}^n k' \left(\left\| \frac{x - x_i^m}{h} \right\|^2 \right) \right] \geq 0$, on obtient, pour chaque échantillon x_k^m avec $k=1..n$:

$$MS_{h,K}(x_k^m) = \frac{nh \hat{\nabla}f_{h,K}(x_k^m)}{2c g_{h,K}(x_k^m)} \quad (74)$$

Cette équation traduit que le terme du mean shift pointe toujours dans la direction du gradient de l'estimée de la densité et donc vers le mode local de l'histogramme lissé.

En utilisant la formule (74), la valeur de l'échantillon x_k^m est mise à jour de façon à le déplacer vers le mode. On pose alors $x_k^{m+1} = x_k^m + MS_{h,K}(x_k^m)$.

La convergence est obtenue à l'itération p lorsque $\forall k \in [1..n], x_k^{p+1} = x_k^p$. C'est à dire qu'il faut que l'estimation du gradient de l'histogramme lissé soit nulle en chaque échantillon, ou encore que chaque échantillon se situe sur le mode local de l'histogramme.

La convergence de l'algorithme mean shift a été démontrée par Cheng pour le noyau normal [Cheng 1995]. De façon plus générale, si K a un profil convexe et strictement décroissant, i.e. $k'(x) < 0, \forall x \in \mathfrak{R}$, la convergence est également assurée [Comaniciu 2002].

4.2.3.4. Les noyaux

Un noyau est donc une forme bilinéaire définie par :

$$K : \begin{cases} \Omega x \Omega \rightarrow \mathfrak{R} \\ (x, y) \rightarrow K(x, y) \end{cases}$$

Deux types de noyaux sont couramment utilisés :

- les noyaux convolutifs : $K(x, y) = K(x - y)$
- les noyaux multiplicatifs : $K(x, y) = K(x/y)$

Le « kernel trick » est introduit dans [Aizerman 1964] et est fondé sur la condition de Mercer spécifiant que toute forme positive, semi définie peut se décomposer en un produit scalaire dans un espace de Hilbert :

$$K(x, y) = \langle \phi(x), \phi(y) \rangle$$

Le produit scalaire est alors associé à une distance et permet de définir une mesure de similarité. Ainsi, en estimant une ddp f par la technique de Parzen, une distance est définie, entre f et sa réalisation. Dans le cas mean shift, les noyaux satisfaisant les conditions de Mercer et de convergence énoncées dans le paragraphe précédent sont nombreux. Les noyaux Normal (K_N de profil k_N) et Epanechnikov (K_E de profil k_E) sont les plus utilisés.

$$K_N(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\|x\|^2}{2}\right) & \text{si } \|x\| \leq 1 \\ 0 & \text{sinon} \end{cases} \quad K_E(x) = \begin{cases} \frac{1}{2}(1 - \|x\|^2) & \text{si } \|x\| \leq 1 \\ 0 & \text{sinon} \end{cases}$$

$$k_N(x) = \begin{cases} \exp\left(-\frac{x}{2}\right) & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases} \quad k_E(x) = \begin{cases} 1 - x & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

Le noyau Epanechnikov est montré être optimal au sens des moindres carrés pour estimer une distribution composée d'un mélange d'un nombre fini de gaussiennes (chapitre 2.7 de [Wand 1994]). Une comparaison des noyaux sur ces distributions est proposée en terme d'efficacité notée e .

Soit N_E le nombre d'échantillons nécessaires à l'estimation d'une telle distribution afin d'obtenir une erreur quadratique moyenne, notée E_{rms} , entre la ddp et son estimée. Pour obtenir la même erreur avec un autre noyau il faudra utiliser N échantillons. L'efficacité e est définie par :

$$e = \frac{N}{N_E}$$

Les efficacités des noyaux Normal et Triangulaire, entre autre, valent respectivement :

$$e_{Normal} = 0.951 \quad e_{Triangulaire} = 0.982$$

C'est à dire qu'il faudra, à erreur quadratique moyenne constante, 5% d'échantillons en plus si on effectue l'estimation d'un mélange de gaussiennes avec le noyau normal au lieu d'utiliser le noyau d'Epanechnikov.

4.2.3.5. Apport du domaine joint spatial / amplitude

L'algorithme du mean shift est adapté au traitement d'image pour tenir compte du voisinage spatial [Comaniciu 2002]. Le noyau est alors un noyau produit décomposé en deux parties : $K(x) = K_S(x) \cdot K_A(x)$ où $K_S(x)$ est le noyau dans le domaine spatial et $K_A(x)$ celui de l'espace des niveaux de gris ou de couleur (espace RGB, LUV...), appelé de façon plus générale espace caractéristique.

L'algorithme est le même que celui présenté au §4.2.3.3 mais il se déroule simultanément dans les domaines spatial et caractéristique. Deux largeurs de bande (h_S dans le domaine spatial et h_A dans le domaine caractéristique) sont alors nécessaires. Pour chaque pixel, et à chaque itération, les pixels utilisés pour estimer la statistique devront être à la fois dans la bande spatiale et la bande caractéristique.

Soient $\{z_1^m, z_2^m, \dots, z_n^m\}$ les données dans l'espace joint à l'itération m .
 $\forall i \in [1 \dots n]$, $z_i^m = \{s_i^m, a_i^m\}$ où s_i^m est la composante spatiale et a_i^m la composante caractéristique de la donnée i à l'itération m .

L'estimée de f et de son gradient s'écrivent alors :

$$\hat{f}(z_j^m) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_A h_S} K_A \left(\frac{a_i^m - a_j^m}{h_A} \right) K_S \left(\frac{s_i^m - s_j^m}{h_S} \right) \quad (75)$$

$$\hat{\nabla} f_{h,K}(z_j^m) \equiv \nabla \hat{f}_{h_A, K_A, h_S, K_S}(z_j^m) = \frac{\partial \hat{f}(z_j^m)}{\partial a_j^m} + \frac{\partial \hat{f}(z_j^m)}{\partial s_j^m}$$

D'après l'équation (75), seuls les pixels satisfaisant les deux conditions suivantes contribueront aux mean shift :

$$\forall i \in [1 \dots n], a_i^m \in [a_j^m - h_A; a_j^m + h_A] \quad (76)$$

$$\forall i \in [1 \dots n], s_i^m \in [s_j^m - h_S; s_j^m + h_S] \quad (77)$$

Soit Ω_j l'ensemble des pixels de l'image vérifiant les conditions des équations (76) et (77). Cet ensemble est représenté sur la figure 66 et correspond au disque gris foncé auquel est ôté la zone gris clair.

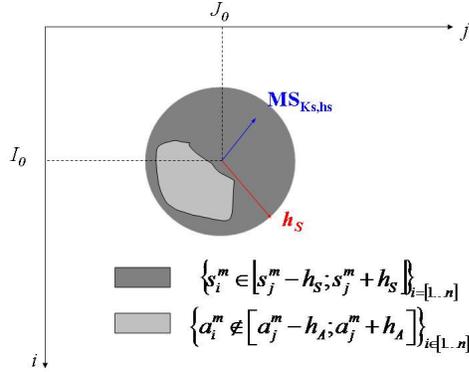


Figure 66 – Illustration du mean shift dans le domaine joint

De la même façon qu'au §4.2.3.3, un terme mean shift pour chaque espace est obtenu :

$$MS_{h_A, K_A}(a_j^m) = \frac{\sum_{i \in \Omega_j} a_i^m \cdot k'_A \left(\left\| \frac{a_j^m - a_i^m}{h_A} \right\|^2 \right)}{\sum_{i \in \Omega_j} k'_A \left(\left\| \frac{a_j^m - a_i^m}{h_A} \right\|^2 \right)} - a_j^m \quad MS_{h_S, K_S}(s_j^m) = \frac{\sum_{i \in \Omega_j} s_i^m \cdot k'_S \left(\left\| \frac{s_j^m - s_i^m}{h_S} \right\|^2 \right)}{\sum_{i \in \Omega_j} k'_S \left(\left\| \frac{s_j^m - s_i^m}{h_S} \right\|^2 \right)} - s_j^m$$

Le terme en mean shift global, utilisé pour la condition d'arrêt, s'écrit alors :

$$MS_{h_A, K_A, h_S, K_S}(z_j^m) = \sqrt{\left[MS_{h_A, K_A}(a_j^m) \right]^2 + \left[MS_{h_S, K_S}(s_j^m) \right]^2}$$

Les valeurs caractéristiques des pixels (niveaux de gris, RGB...) se déplacent comme précédemment dans le domaine caractéristique vers le mode le plus proche de l'histogramme lissé. De la même façon, les coordonnées spatiales des pixels sont également mises à jour dans le domaine spatial. Les pixels de l'image sont alors déplacés et se regroupent dans des bassins d'attractions qui sont l'équivalent des modes dans le domaine caractéristique. A

chaque itération, un nombre d'échantillons croissant autour de ces bassins est considéré pour effectuer un meilleur lissage d'histogramme et donc une meilleure estimation du mode de l'histogramme dans l'espace caractéristique. A la fin du filtrage, les pixels sont assignés à leurs positions initiales et leurs valeurs caractéristiques sont celles obtenues après convergence de l'algorithme mean shift. Effectuer le filtrage dans le domaine joint permet ainsi de tenir compte du voisinage spatial, porteur d'information.

4.2.3.6. Lien entre Mean shift et SVM

Le support vector machine (SVM) est une méthode de classification basée, comme le mean shift, sur les méthodes à noyaux. Elle a montré de bonnes performances dans de nombreux domaines du traitement d'images (classification de textures, d'expressions faciales, de textes... [Clopinet 2006]). C'est pourquoi il est intéressant d'effectuer un rapprochement de ces deux algorithmes.

Pour comprendre intuitivement le principe du SVM, un cas simple à deux classes (échantillons rouges et bleus) et en dimension 2 est utilisé (cf. figure 67). L'idée du SVM est de trouver le plan P optimal permettant de séparer ces deux classes. Dans ce cadre, l'optimalité consiste à maximiser la marge entre le plan séparateur et les vecteurs supports, c'est-à-dire les vecteurs de chaque classe les plus proches du plan séparateur.

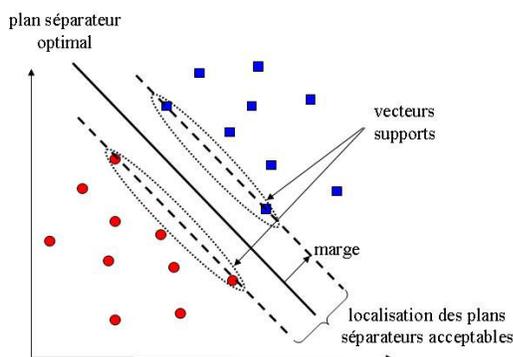


Figure 67 – Principe du SVM pour deux classes linéairement séparables

Lorsque les classes sont linéairement séparables (cf. figure 67), intuitivement, cette séparation semble faisable. Néanmoins il y a des cas où les deux classes ne sont pas linéairement séparables (cf. figure 68a).

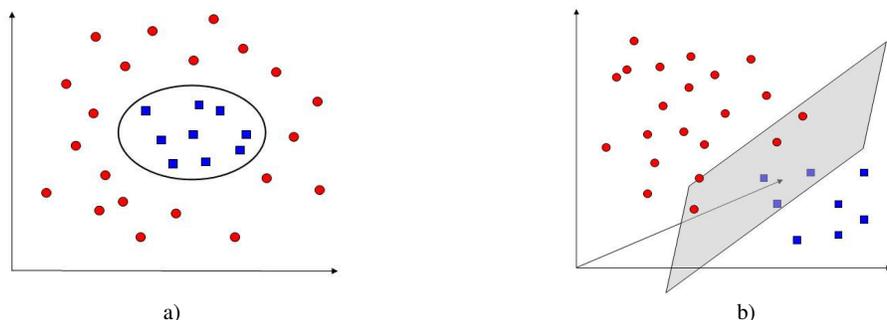


Figure 68 – a) SVM pour deux classes non linéairement séparables, b) espace de re-description où les classes sont linéairement séparables

L'idée du SVM est de projeter les données dans un autre espace, appelé espace de re-description, où elles sont linéairement séparables (cf. figure 68b). Cette projection est

effectuée grâce à un noyau satisfaisant les mêmes contraintes que celles exposées dans le cas du mean shift (cf. §4.2.3.4).

Dans la suite de ce paragraphe, il est montré que le mean shift peut être considéré comme un cas particuliers de SVM. Considérons maintenant l’algorithme mean shift, sans extension au domaine spatial, appliqué au cas de la figure 68a. L’espace initial est un espace mono-dimensionnel qui correspond aux valeurs d’amplitude. Lorsque le lissage de l’histogramme global est effectué par technique de Parzen, cela revient à projeter l’ensemble des données dans un espace bidimensionnel. Cet espace de re-description est l’espace dans lequel est représentée la densité de probabilité associée aux échantillons d’amplitude. Le lissage d’histogramme apparaît alors comme un cas particuliers de projection où le choix de l’espace de re-description est fixé par la méthode. La séparation des classes, qui sont dans ce cas les modes des statistiques, peut être alors effectuée dans le domaine de re-description de façon linéaire par un seuillage (cf. figure 69).

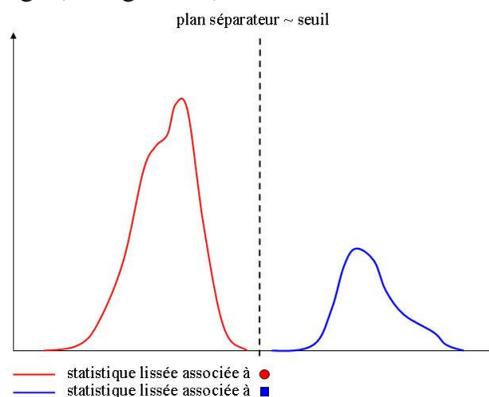


Figure 69 – Rapprochement mean shift / SVM

Lorsque l’algorithme mean shift est étendu au domaine spatial, le principe est identique sauf que l’espace initial des données est de dimension 3 (localisation, amplitude). L’espace de re-description est alors l’espace des statistiques locales de dimension 4. Une fois encore, la séparation de deux classes dans cette espace de re-description peut être effectué par un seuillage sur les valeurs d’amplitude.

Il faut également ajouter que le SVM est un algorithme supervisé, c’est-à-dire qu’il a besoin d’un apprentissage sur des données dont les classes sont connues pour pouvoir déterminer les paramètres du plan séparateur. L’algorithme mean shift est un algorithme non-supervisé. Initialement, les modes des statistiques ne sont pas connus mais ils sont estimés de façon itérative par l’algorithme.

4.2.3.7. Conclusion

Le principe de fonctionnement de l’algorithme mean shift développé par Fukunaga et Hostetler a été détaillé. Cet algorithme est itératif et peut se décomposer en deux parties. Tout d’abord un lissage par méthode à noyaux de l’histogramme est effectué. Ensuite la détermination de la position d’un mode local de cet histogramme lissé permet de déplacer les échantillons vers ce mode. Cet algorithme a ensuite été étendu au traitement d’image par Comaniciu et Meer en tenant compte du voisinage spatial. Il faut remarquer que le même algorithme de base est utilisé conjointement dans le domaine caractéristique et dans le domaine spatial. Sur le même principe, nous proposons une extension de cet algorithme au §4.3.

4.2.4. Comparaison du filtre mean shift aux filtres paramétriques en imagerie RSO

4.2.4.1. Introduction

Dans ce paragraphe nous nous intéressons au comportement du filtre mean shift de Commaniciu et Meer en imagerie RSO haute résolution. Nous exposons dans un premier temps les limitations au bon fonctionnement du filtre mean shift sur ces images à cause des statistiques des images RSO en amplitude. Ensuite, une méthode d'évaluation et de comparaison objective des filtres, paramétriques et non-paramétriques, est présentée. Les divers filtres sont comparés sur un jeu d'images synthétiques et sur images RAMSES haute résolution en bande X.

4.2.4.2. Difficultés liées à l'imagerie RSO

Le filtrage des images RADAR constitue une première étape vers l'extraction de primitives ou une éventuelle segmentation des images. Dans l'annexe B.4, les lois statistiques présentes sur les images de Dunkerque sont validées. Les statistiques des images en amplitude ne sont pas gaussiennes mais du type Rayleigh-Nakagami pour les ombres (lorsqu'elles ne sont pas bruitées par les échos triples) et pour le sol. Notre étude statistique montre que le modèle empirique de Fisher s'adapte également bien à nos données pour tout type de zones. Les limites des estimations paramétriques utilisant un modèle apparaissent lorsqu'il y a un mélange de lois dans une même zone. Pour les images en amplitude, ce mélange intervient par la présence de réflexions triples dans les ombres (cf. §2.4.4.3). En monovue ($L=1$), un mélange de lois Rayleigh-Nakagami ne présente qu'un seul mode pour plusieurs contrastes (cf. figure 70).

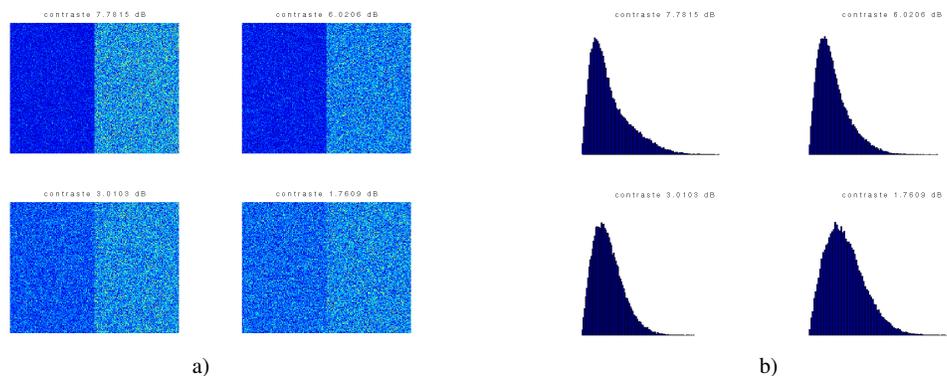


Figure 70 – a) images synthétiques comportant deux zones suivant chacune une loi Rayleigh Généralisée pour les contrastes (de gauche à droite et de haut en bas) 7.78dB, 6.02dB, 3.01dB et 1.78dB, b) histogrammes correspondants

En fait, chaque zone présente un mode mais ces modes ne sont pas distincts à cause de la nature des lois Rayleigh-Nakagami. L'application directe sur les images RSO du mean shift et des autres filtres semble donc difficile.

4.2.4.3. Méthode d'évaluation

4.2.4.3.1. Présentation des images synthétiques

Chaque image synthétique est composée de deux zones de 128x256 pixels. Ces zones ont été générées aléatoirement sous Matlab et suivent une loi Rayleigh-Nakagami généralisée monovue ($L=1$) et multivue ($L=4$) de paramètre μ . Le rapport μ_1/μ_2 définit le contraste entre les deux zones de l'image. Les contrastes testés vont de 0.41dB à 7.78dB. Les tests ont été

effectués sur un jeu de 100 images par contraste. Pour valider nos simulations, nous représentons ci-dessous les distributions théoriques et empiriques des mélanges de lois Rayleigh-Nakagami présentes sur les images.

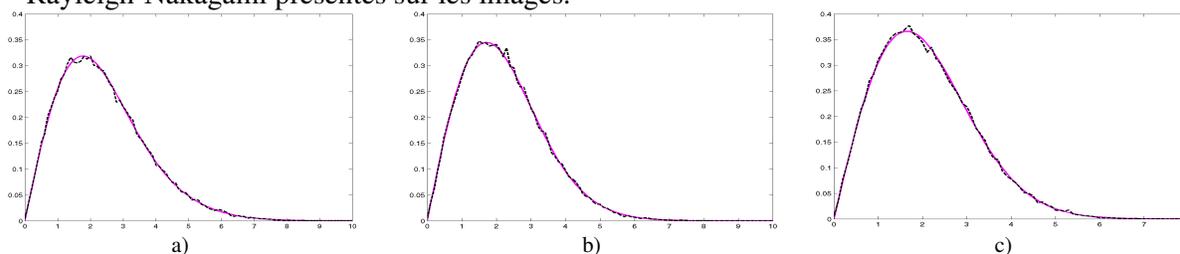


Figure 71 – distributions théoriques (trait plein) et empiriques (trait discontinu) du mélanges de deux loi de Rayleigh-Nakagami monovues pour des contrastes de a) 3.01dB, b) 1.78dB, c) 0.79dB

4.2.4.3.2. courbes COR théoriques

Les courbes COR (Courbes d’Opération en Réception) représentent la probabilité de détection en fonction de la probabilité de fausse alarme d’un algorithme pour un jeu de paramètre donné. Pour tracer une courbe COR, il faut donc plusieurs couples de probabilité que l’on peut obtenir par différents seuillages.

Prenons l’exemple de deux gaussiennes p_1 et p_2 :

$$p_1(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right)$$

$$p_2(x) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(x-\mu_2)^2}{2\sigma_2^2}\right)$$

Les probabilités de détection et de fausse alarme pour un seuil donné sont représentées sur la figure 72.

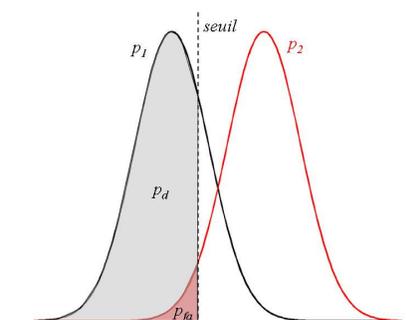


Figure 72 – Seuillage pour l’obtention d’un couple $\{p_d, p_{fa}\}$ pour tracer une courbe COR

Les expressions théoriques des probabilités de détection et de fausse alarme en fonction du seuil sont :

$$p_d(x) = \int_{-\infty}^{\text{seuil}} p_1(t) dt = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\text{seuil} - \mu_1}{\sqrt{2}\sigma_1} \right) \right) \quad (78)$$

$$p_{fa}(x) = \int_{-\infty}^{\text{seuil}} p_2(t) dt = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\text{seuil} - \mu_2}{\sqrt{2}\sigma_2} \right) \right) \quad (79)$$

où $\operatorname{erf}(x)$ est définie par :

$$\operatorname{erf}(x) = \int_0^x \frac{2}{\sqrt{\pi}} \exp(-u^2) du$$

Les courbes COR théoriques pour différentes gaussiennes peuvent alors être tracées :

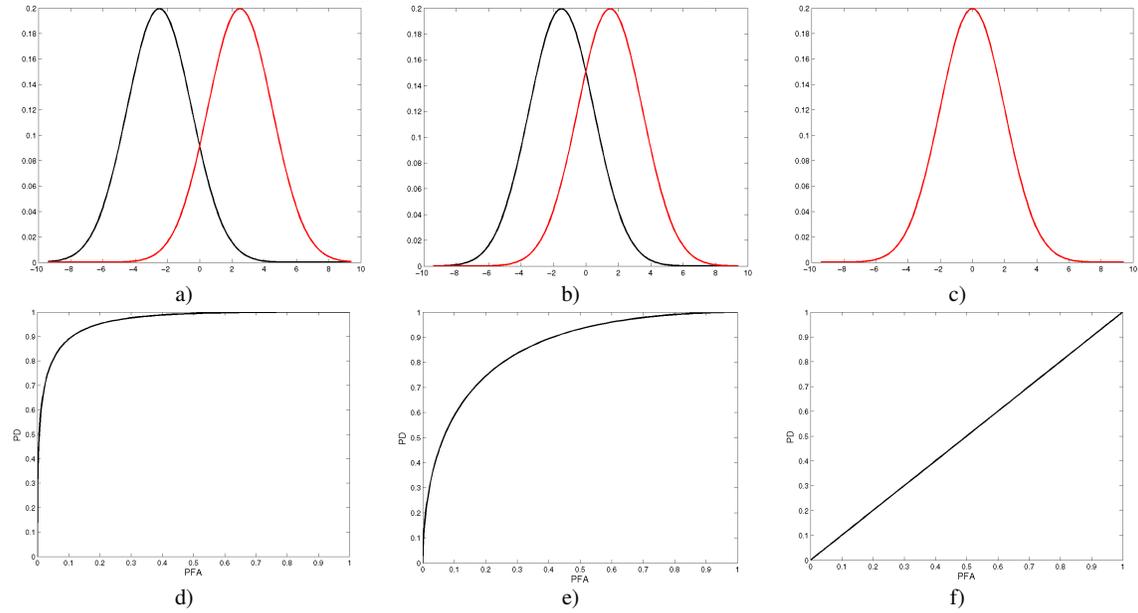


Figure 73 – deux gaussiennes de moyenne a) 2.5 et -2.5, b) 1.5 et -1.5, c) 0 et d'écart-type 2
d), e) et f) courbes COR correspondantes

Plus les moyennes des gaussiennes sont proches et plus la séparation des deux gaussiennes par un seuil sera mauvaise, ce qui se traduit sur les courbes COR théorique par un aplatissement (cf. figure 73b). Dans le cas extrême où les deux gaussiennes sont identiques (cf. figure 73c), les équations (78) et (79) donnent $p_d = p_{fa}$.

Pour deux distributions de Rayleigh, l'expression théorique des courbes COR n'est pas facilement calculable. Nous avons tracé sur la figure 74 ces ddp pour deux réalisations sur un nombre élevé d'échantillons (10000).

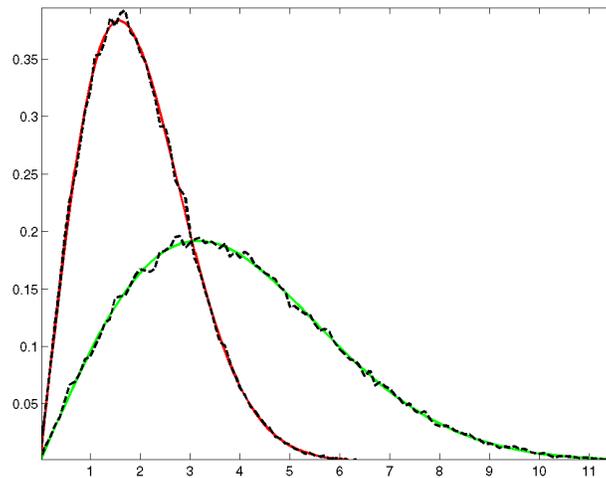


Figure 74 – Lois de Rayleigh théoriques monovue pour un contraste de 7dB et leurs estimées en pointillés

L'erreur quadratique moyenne entre la densité théorique et la densité obtenue par simulation est de l'ordre $7.24 \cdot 10^{-4}$. Les courbes COR théoriques peuvent donc être approchées par une méthode numérique.

Sur la figure ci-dessous les courbes COR théoriques sont représentées dans le cas monovue :

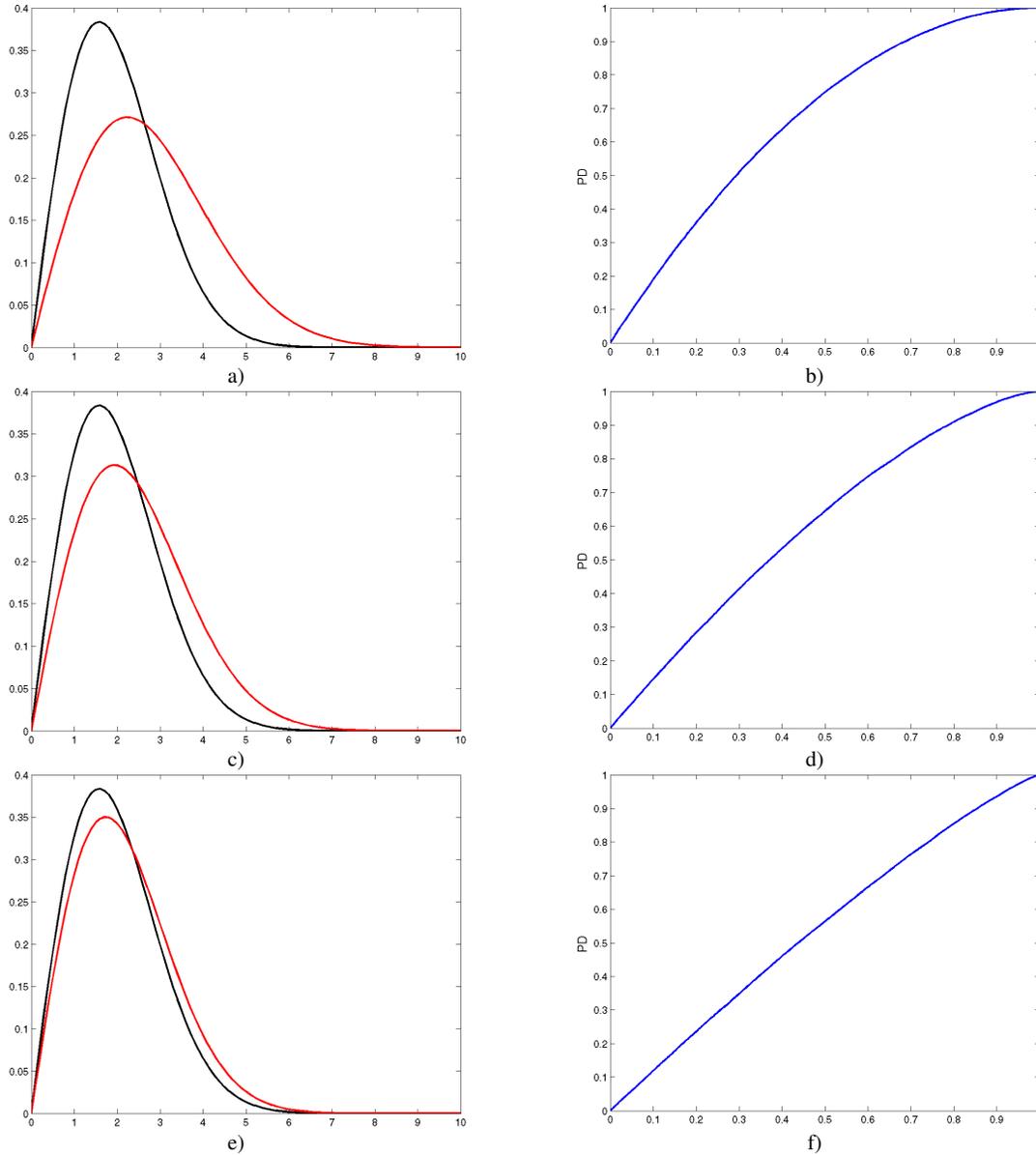


Figure 75 –ddp pour un contraste de a) 3.01dB, c) 1.76dB et e) 0.78dB, courbes COR correspondantes b), d) et f)

Contrairement au cas gaussien, les courbes COR ne sont plus symétriques car la loi de Rayleigh-Nakagami ne l'est pas. Pour un contraste de 3dB, la courbe COR n'est pas bonne car il y a une zone de recouvrement importante des deux ddp.

Pour les mêmes contrastes, en multivue ($L=4$), les courbes COR sont représentées en figure 76. En multivue, les lois Rayleigh généralisée tendent à devenir symétriques. Par conséquent, les courbes COR deviennent également symétriques. Le passage en multivue implique un recouvrement plus faible des deux ddp d'où une amélioration des courbes COR par rapport au monovue (figure 75a et de la figure 76a). Pour des contrastes plus forts (à partir de 7dB environ), les ddp peuvent même être complètement séparées. Pour des contrastes très faibles (0.78dB), le passage en multivue apporte seulement une légère amélioration.

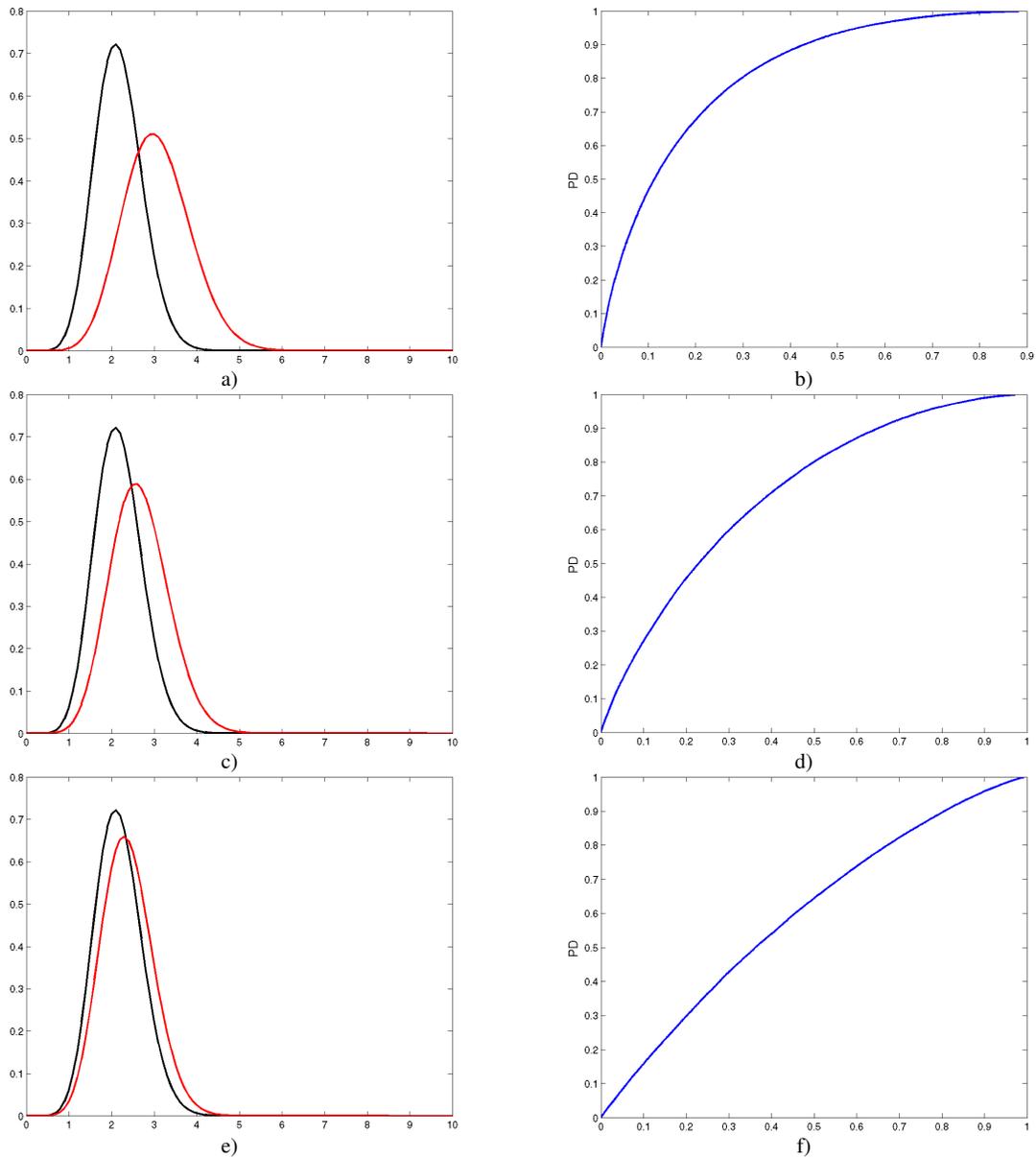


Figure 76 – ddp pour un contraste de a) 3.01dB, c) 1.76dB, et e) 0.78dB, courbes COR correspondantes b), d) et f)

Nous voulons effectuer une comparaison objective des différents filtres. L'utilisation en post-traitement de méthodes de classification (kmeans ou autres) ayant une influence sur les résultats, les courbes COR sont choisies pour effectuer cette comparaison.

Jusqu'à présent, les courbes COR ont été calculées à partir d'un histogramme. Pour calculer une courbe COR sur une image, il suffit de définir la zone à détecter (cf. figure 77).

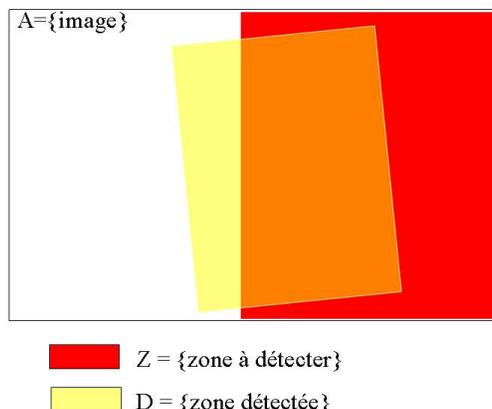


Figure 77 – Représentation schématique d'une image A, de la zone à détecter Z et de la zone détectée par seuillage D

Sur les images synthétiques (notée A), deux classes (sombre, claire) sont présentes. La zone de droite est choisie en tant que zone à détecter (Z). En effectuant, après filtrage de l'image, un seuillage à une valeur arbitraire, une zone détectée, notée D, est obtenue. Par comparaison de la zone détectée et de Z, un couple de valeurs {probabilité de détection (p_d) ; probabilité de fausse alarme (p_{fa})} est obtenu.

$$\text{Ces probabilités sont définies par : } p_d = \frac{D \cap Z}{Z} \text{ et } p_{fa} = \frac{D - D \cap Z}{A - Z} .$$

L'intérêt de cette méthode d'évaluation est d'être totalement indépendante des filtres comparés.

4.2.5. Performances sur images synthétiques

4.2.5.1. Introduction

Cette comparaison s'effectue sur des images synthétiques présentant des lois de Rayleigh-Nakagami généralisée. On se place dans une situation idéale pour les filtres paramétriques car l'hypothèse portant sur la nature des lois statistiques au sein de ces images est vérifiée. Nous montrons dans ce paragraphe que le mean shift, sans faire d'hypothèses sur le type de loi présent sur les images synthétiques, obtient des résultats similaires aux filtres paramétriques.

Dans un premier temps, l'influence de la frontière sur ces images tests est analysée. Le filtre mean shift est ensuite comparé au filtre médian, non paramétrique, et très utilisé pour débruiter les images. L'influence d'un préfiltrage multivue sur les performances du mean shift est également étudiée. Puis nous nous intéressons à l'influence des paramètres de bandes du mean shift sur la qualité du filtrage. Enfin nous comparons le filtre mean shift aux filtres paramétriques usuels en imagerie RSO.

4.2.5.2. Performances du mean shift sur zones homogènes

Pour divers contrastes, les performances du mean shift sont comparées sur les images synthétiques entières (256x256 pixels) et ces mêmes images autour de la frontière entre les deux zones (16x256 pixels). Cela permet d'analyser l'influence de la frontière sur les performances du filtrage. Les paramètres de bande amplitude et spatiale pour les images monovues ont été choisis arbitrairement : $h_A=2$ et $h_S=5$ (ce qui correspond à un voisinage 9x9). L'influence de ces paramètres sur la qualité du filtrage sera étudiée au §4.2.5.4.

Lorsque aucun filtrage n'est utilisé, les performances sur les images entières et autour des contours sont les mêmes (cf figure 78), ce qui est normal car les distributions à l'intérieur des deux demi-bandes autour de la frontière n'ont aucune raison d'être différentes de celles à l'intérieur des deux zones. En revanche, lorsque le filtre mean shift est utilisé, il est visible que les performances sont moins bonnes lorsqu'on ne considère que la frontière. En effet, autour de la zone frontière, à l'intérieur de la fenêtre de filtrage, il y a un mélange de statistiques et estimer le mode devient plus difficile.

La bande considérée autour de la frontière entre les deux zones est faible. Des tests ont montrés que, pour un contraste de 3dB et une p_{fa} de 5%, la p_d passe de 96% à 94.5% lorsque la frontière est considérée. Plus le contraste diminue plus cette différence est négligeable car les pixels des deux zones suivent des lois de plus en plus similaires et la différence entre les deux zones s'estompe.

Dans la suite de ce paragraphe, l'influence de la frontière est donc considérée comme négligeable pour déterminer la capacité du mean shift à différencier deux zones homogènes, de statistiques différentes, par leur modes.

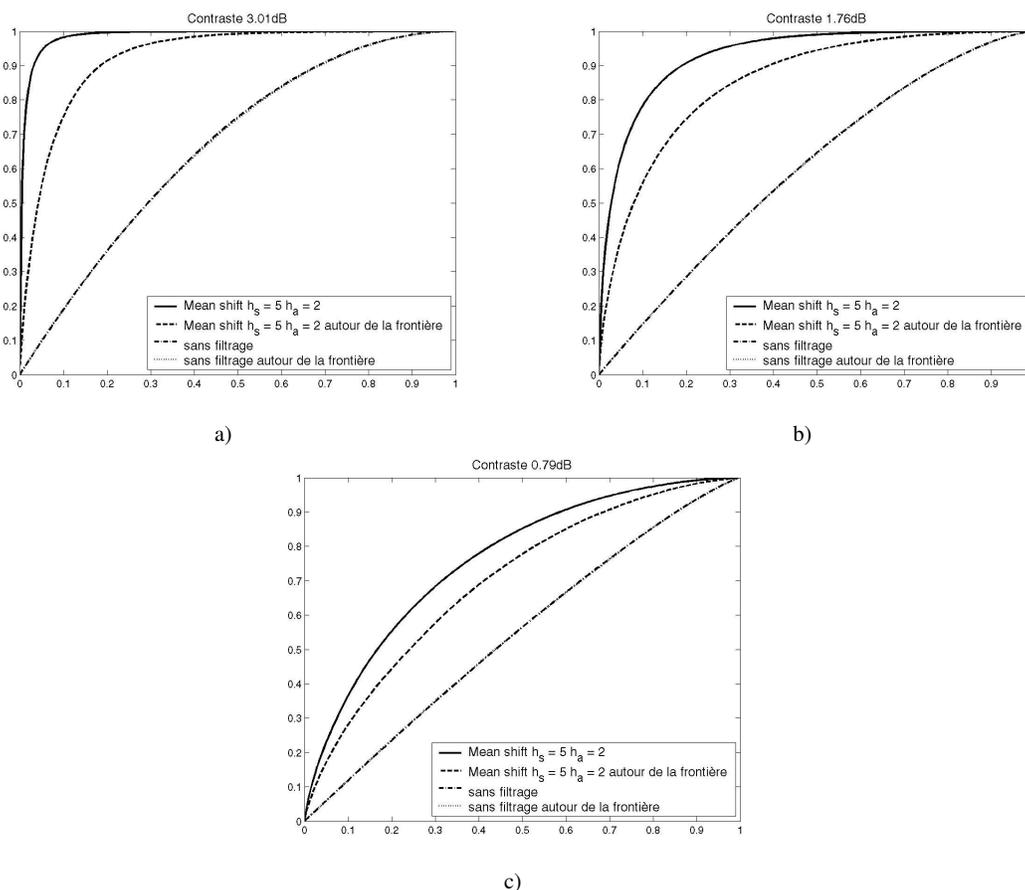


Figure 78 – Influence du contour sur les performances du mean shift pour des contrastes de a) 3.01dB, b) 1.76dB et c) 0.79dB

4.2.5.3. Influence du préfiltrage

Le multivue est un préfiltrage usuel en imagerie RSO. Il permet de réduire le chatoiement au détriment de la résolution. Pour observer l'influence du multivue sur le résultat du filtrage, le mean shift doit être appliqué sur une zone identique en simple et multivue.

Choisir une largeur de bande spatiale h_s revient à utiliser, lors du filtrage, les pixels inclus dans un disque de diamètre $2h_s-1$. Un multivue est effectué en bloc carré de L pixels, le diamètre du disque utilisé sur l'image multivue doit être $(2h_s-1)L^{-1/2}$. Choisir une bande spatiale de 5 en simple vue revient à prendre une bande spatiale de 2,75 en multivue 2×2 ($L=4$).

Sur la figure 79 sont représentées, pour une image synthétique de 1,76dB, les courbes COR sans filtrage et avec filtrage mean shift en simple et multivue. Pour un même « voisinage » spatial, le préfiltrage multivue améliore les performances du mean shift en réduisant le chatoiement. C'est pourquoi, dans la suite de ce chapitre, seuls les résultats sur images multivues sont analysés. Les résultats sur images monovues sont présentés en annexe D.

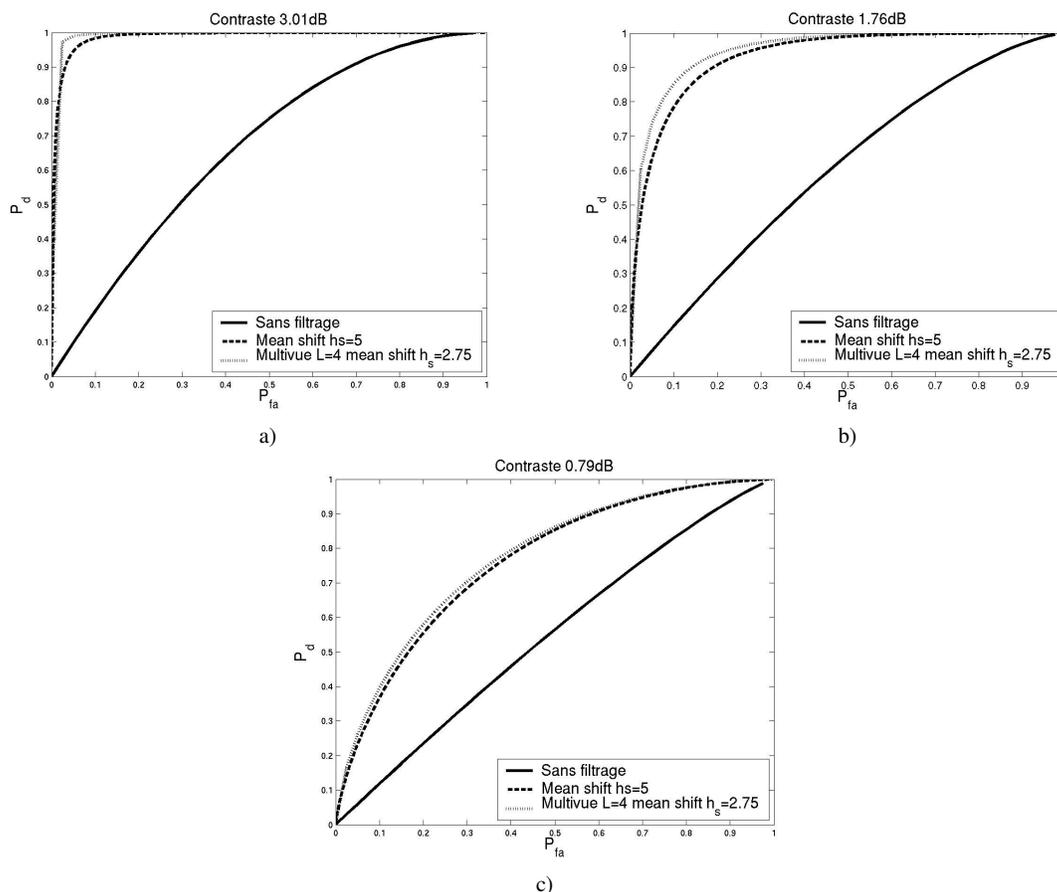


Figure 79 – Influence du préfiltrage sur le mean shift pour des contrastes de a) 3.01dB, b) 1.76dB et c) 0.79dB

4.2.5.4. Influence des paramètres de bandes spatiale et amplitude

L'influence du paramètre de bande amplitude (h_A) sur les performances du filtrage est représentée sur la figure 80 dans le cas d'images multivues 2x2. Sur chaque série d'image, la moyenne de la zone de plus faible amplitude est de 2.23 environ. Les moyennes des zones de plus forte amplitude pour des contrastes de 3dB, 1.76dB et 0.79dB sont respectivement 3.15, 2.73 et 2.44. Les bandes amplitudes testées sont de l'ordre de grandeur de la différence de moyenne entre les deux zones.

Augmenter la bande amplitude permet de considérer des pixels du voisinage spatial dont l'amplitude est plus éloignée de l'amplitude du pixel au centre de la fenêtre de filtrage. Par conséquent, le nombre de pixels utilisés pour l'estimation augmente et le filtrage aboutit à de meilleurs résultats pour tous les contrastes. Cette amélioration des courbes COR s'arrête au moment où tous les pixels sont considérés. Pour $h_A > 2$, les courbes sont pratiquement identiques. Elles n'ont pas été représentées pour ne pas surcharger les figures.

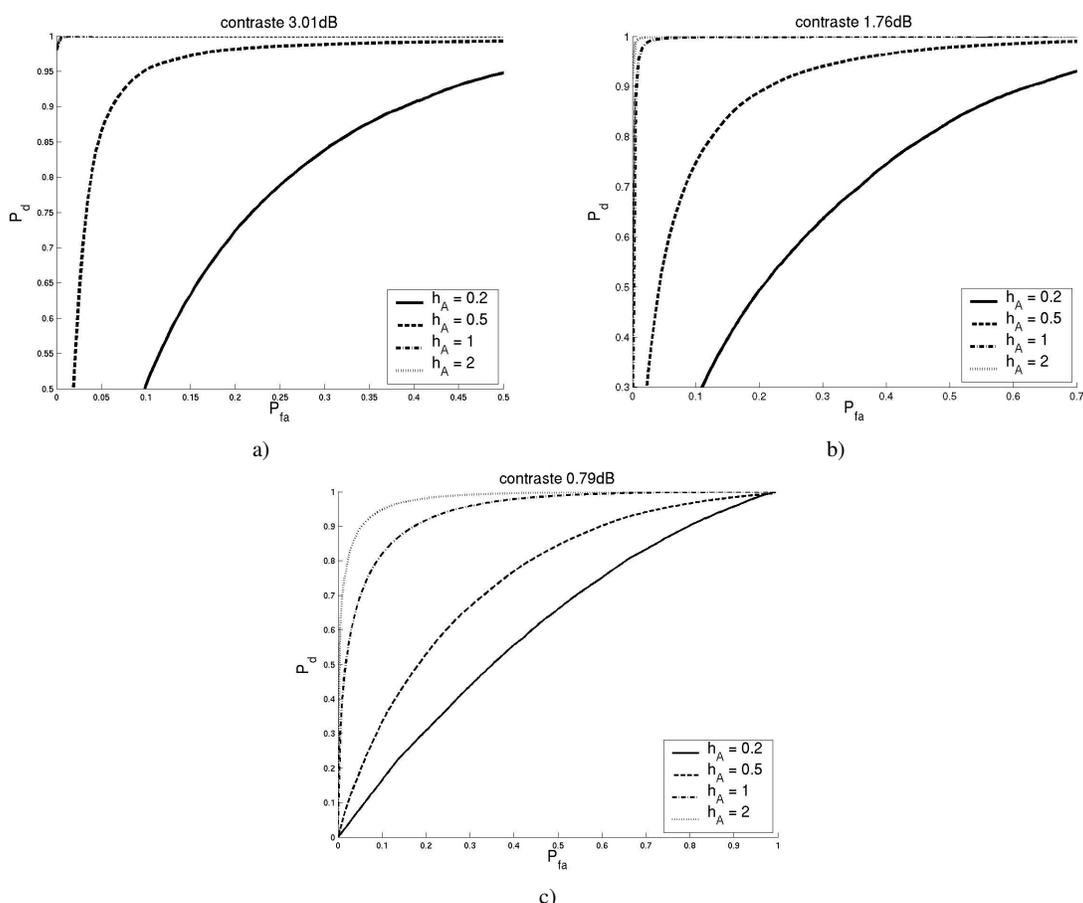


Figure 80 – Influence du paramètre de bande amplitude sur le filtrage mean shift sur images synthétiques multivues ($L=4$) pour des contrastes de a) 3.01dB, b) 1.76dB et c) 0.79dB

Néanmoins, l'utilisation d'une bande amplitude trop grande a aussi pour conséquence de rendre les contours flous (cf. figure 81). L'influence de la largeur de bande amplitude sur le filtrage est visible. Les paramètres de bande amplitude adaptés semblent être 0.75 et 1. Pour un paramètre de bande de 2, la zone frontière est floue.

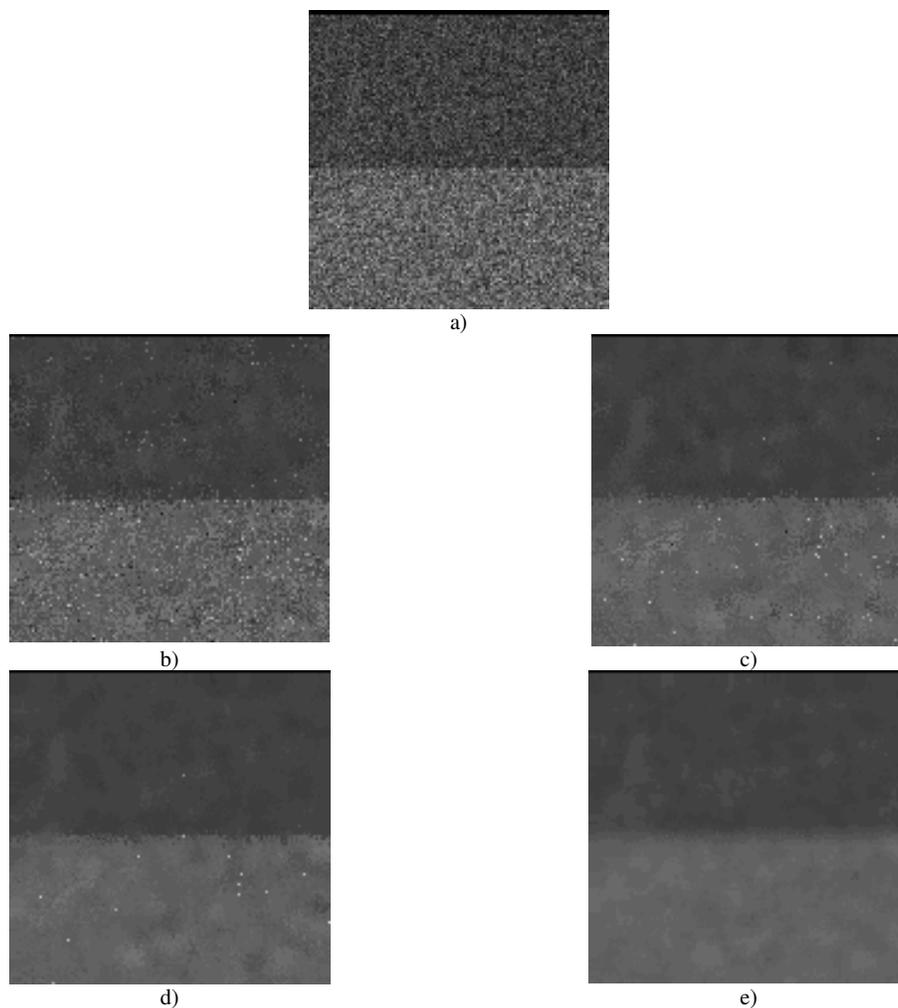


Figure 81 – Influence du paramètre de bande amplitude sur la netteté de la frontière sur une image synthétique multivue ($L=4$) de contraste 3dB, a) image non-filtrée, b) $h_A = 0.5$, c) $h_A = 0.75$, d) $h_A = 1$, e) $h_A = 2$

L'utilisation d'une largeur de bande spatiale (h_s) plus importante (cf. figure 82) permet de prendre en compte un nombre plus important de points et permet d'améliorer l'estimation du mode local. Ceci est visible pour les images de contrastes 0.78dB où un grand nombre d'échantillons est nécessaire pour effectuer la distinction entre les deux régions. Pour des contrastes plus forts (3.01dB et 1.76dB), les courbes COR pour les paramètres de bande spatiale 5 et 7 sont quasiment confondues car le nombre d'échantillons requis pour séparer les modes est déjà atteint.

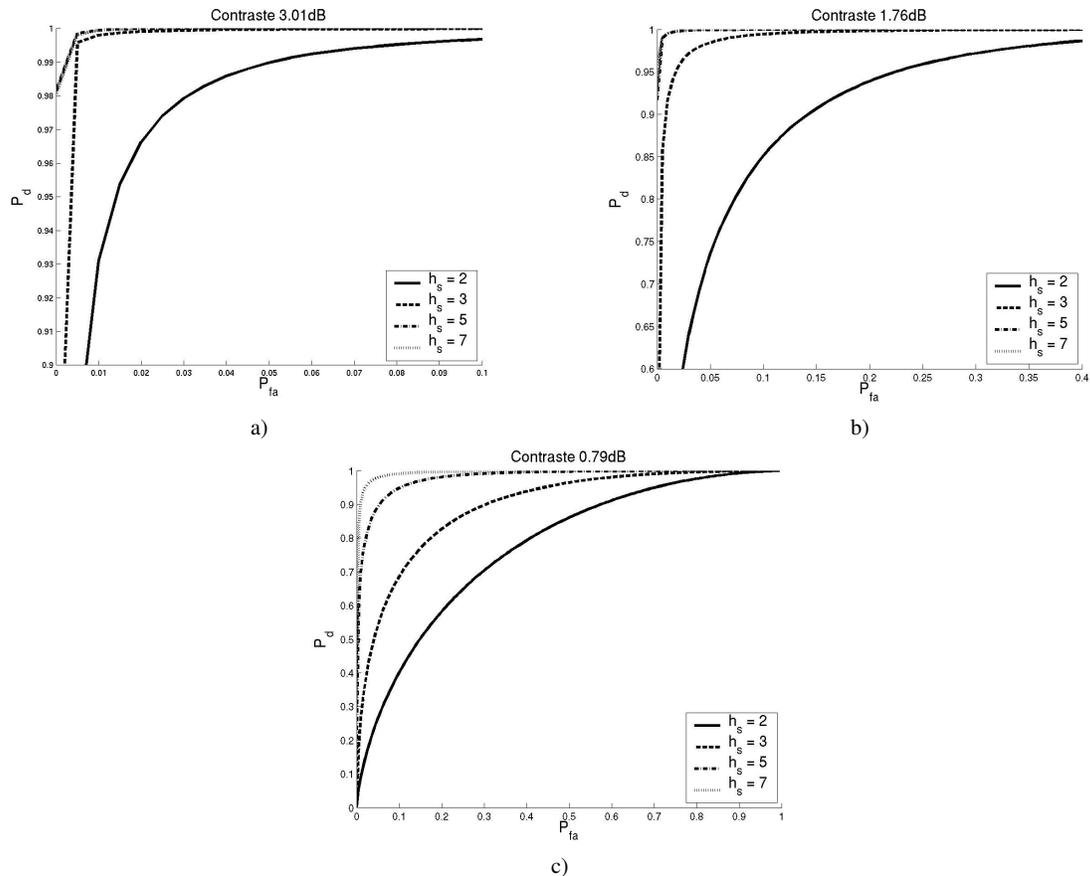


Figure 82 – Influence du paramètre de bande spatiale sur le filtrage mean shift sans préfiltrage sur images synthétiques multivues ($L=4$) pour des contrastes de a) 3.01dB, b) 1.76dB et c) 0.79dB

Il faut noter que pour des fenêtres plus grandes, les filtres sont souvent sanctionnés car le risque d'estimer un mélange de loi au lieu d'une seule est plus grand. Dans le cas du mean shift, lorsque les zones de part et d'autre de la frontière sont très contrastés, le paramètre de bande amplitude permettra d'éliminer certains pixels aberrants.

Lorsque la bande spatiale augmente, l'accroissement des performances se fait au détriment du temps de calcul. Il faut donc effectuer un compromis performance / temps d'exécution.

4.2.5.5. Comparaison de filtres non-paramétriques

Les courbes COR, associées aux filtres mean shift et médian, sont comparées sur les images synthétiques multivues lorsque les deux filtres sont appliqués sur un voisinage 9×9 (figure 83).

Pour de forts contrastes, ces deux filtres aboutissent à de bons résultats et il est difficile de discerner les courbes COR. Pour un faible contraste (figure 83c), le mean shift présente de meilleure performance.

Il faut également préciser que l'utilisation du filtre médian sur un voisinage 9×9 présente plusieurs inconvénients. Tout d'abord, le filtre médian ne préserve pas les marches et rampes rectilignes d'amplitude de taille inférieure à la moitié de la taille de la fenêtre utilisée. Ainsi pour un filtrage médian 9×9 , un bloc isolé de 4×4 pixels au milieu d'une zone homogène, sera effacé, ce qui constitue une perte d'information inacceptable. De plus, le filtre médian érode les arêtes courbes. Ainsi le passage d'un filtre médian $(2n+1) \times (2n+1)$ sur un

disque diminue son rayon de n pixels. Les angles droits sont également érodés. Le choix d'une image adaptée nous a permis de mettre en évidence ces limitations et les résultats sont présentés en annexe D.

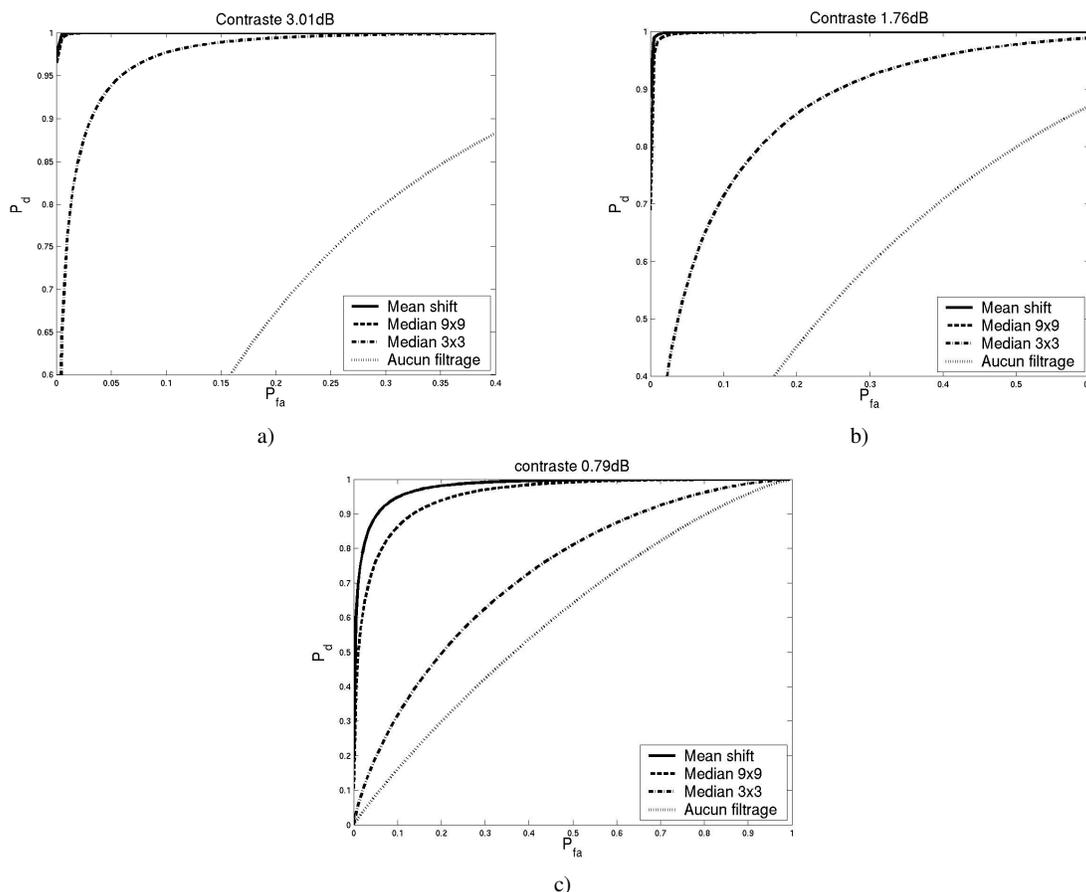


Figure 83 – Performances des filtres non-paramétriques médian 9x9, mean shift et médian 3x3 sur images synthétiques multivues ($L=4$) pour des contrastes a) 3.01dB, b) 1.76dB et c) 0.79dB

Pour limiter les pertes d'informations, le filtre médian est également testé sur un voisinage 3x3. Sur ce voisinage, il élimine seulement le bruit impulsif. L'érosion des arêtes courbes est toujours présente mais la perte d'information est beaucoup moins importante (1 pixel au lieu de 16). Les résultats du filtrage sont en revanche nettement moins bons que ceux du mean shift.

Le filtre mean shift est donc mieux adapté que le filtre médian pour le filtrage des images synthétiques de distribution statistique Rayleigh-Nakagami.

Ces tests ont permis de mettre en évidence une limitation de la méthode d'évaluation choisie : la courbe COR d'un filtre peut être proche de celle d'un autre mais entraîner au passage une perte d'information importante. Dans le cas du filtre médian 9x9, il s'agit d'une perte liée à la suppression d'un bloc de pixels (jusqu'à 4x4).

4.2.5.6. Comparaison des filtres paramétriques

Sur la figure 84, sont représentées les courbes COR des filtres paramétriques et du filtre mean shift sur images multivues ($L=4$). Seuls les meilleurs filtres MEQM et MAP sont représentés. Dans l'étude effectuée, les filtres Frost amélioré et GammaMAP obtiennent les

meilleures performances. Les performances des autres filtres paramétriques en simple et multivue pour des p_{fa} de 1%, 2% et 5% sont fournies en annexe D.

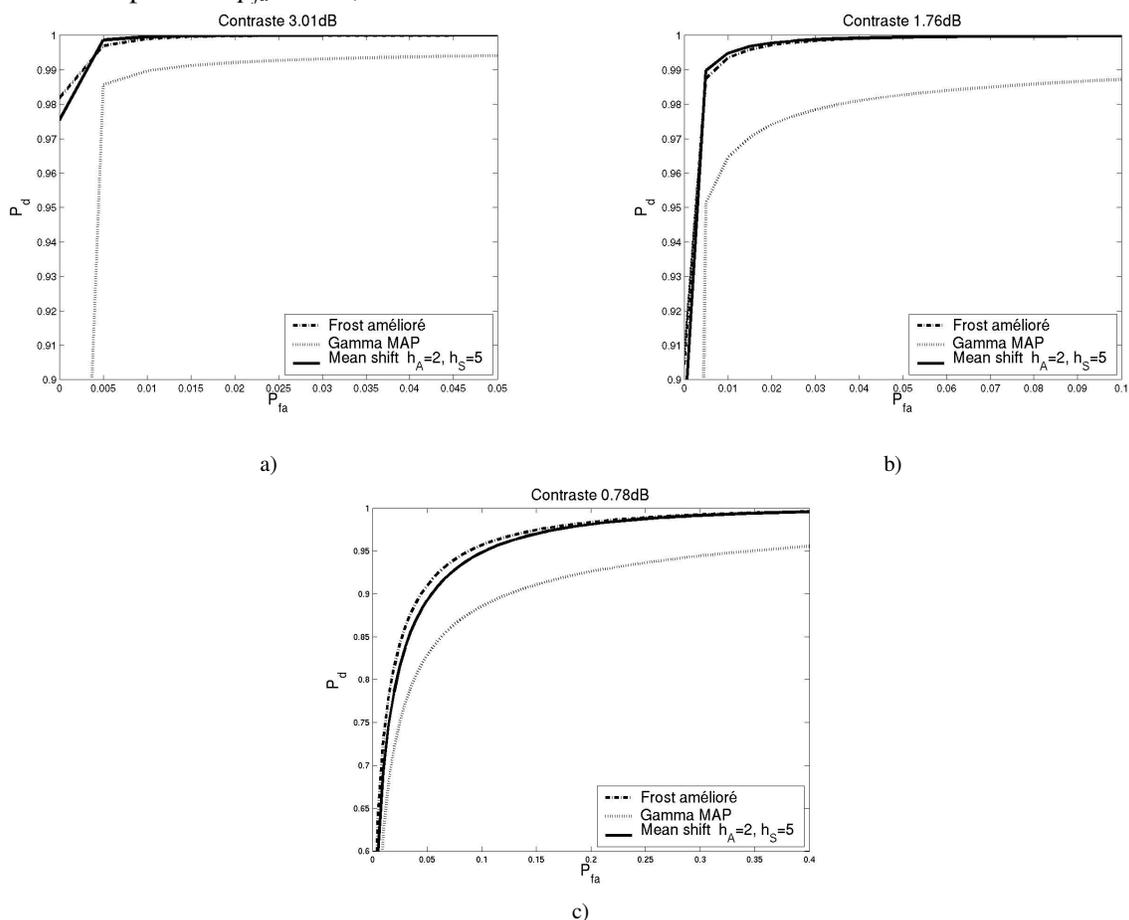


Figure 84 – Comparaison des filtres paramétriques Frost amélioré et GammaMAP et du mean shift sur images synthétiques multivues ($L=4$) pour des contrastes de a) 3.01dB, b) 1.76dB et c) 0.78dB

Les filtres mean shift et Frost fournissent des performances similaires et meilleures que les autres filtres. Néanmoins, ces filtres paramétriques sont utilisés dans le cas idéal car les images synthétiques présentent du speckle complètement développé.

Le filtre mean shift, sans faire d'hypothèses sur la nature des lois au sein des images, fournit des résultats très proches du meilleur filtre paramétrique sur images synthétiques. Le paragraphe suivant analyse les performances de ces filtres sur images RAMSES.

4.2.6. Performances sur images RAMSES

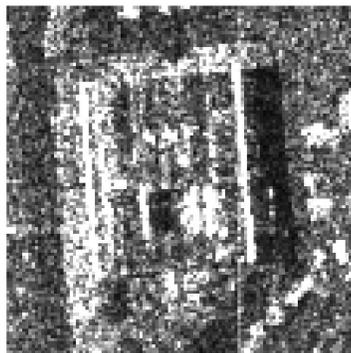
Sur image RAMSES multivue (2×2), les performances du mean shift et des filtres paramétriques sont évaluées pour la détection des ombres. Ces régions ont été délimitées manuellement en utilisant une vue optique de la zone. Cette segmentation manuelle n'est donc pas optimale car les images sont fortement bruitées par des échos triples. Néanmoins elle nous permet de comparer les divers algorithmes.

Pour les filtres paramétriques, des fenêtres 7×7 et 9×9 ont été utilisées et c'est cette dernière qui fournit les meilleures performances. Pour chaque image, les courbes COR du filtre mean shift et des meilleurs filtres paramétriques de type MEQM et MAP sont représentées. Les images résultant d'un filtrage mean shift et du meilleur filtre paramétrique

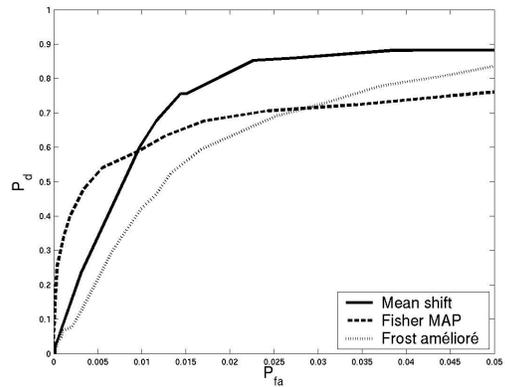
sont également présentées. Les images ont des moyennes et des écart-types du même ordre ce qui nous permet de les comparer pour un même étalement.

Les ombres détectées sont représentées en noir lorsque les images filtrées sont seuillées pour une p_{fa} de 2%. Les zones en gris clair correspondent aux ombres non-détectées. Enfin, en blanc, sont représentées les fausses alarmes.

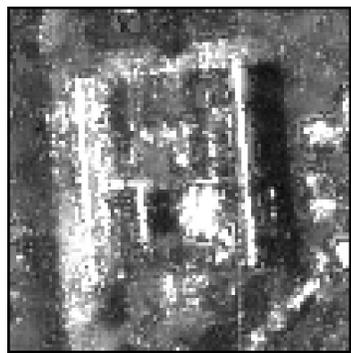
Image 1 : il s'agit d'une imagette 220x220 où les ombres apparaissent relativement bien et sont de formes rectangulaires. Des échos multiples apparaissent dans les ombres.



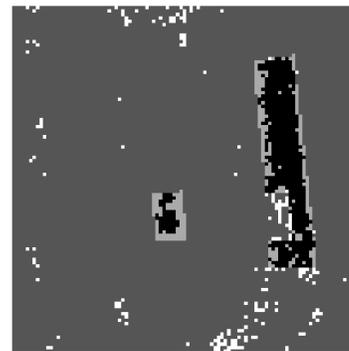
a)



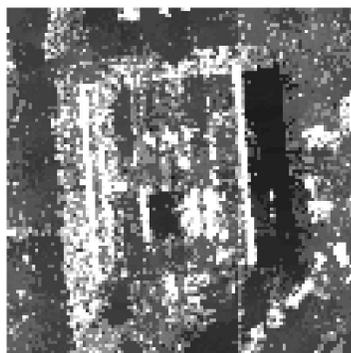
b)



c)



d)



e)



f)

Figure 85 – a) image 1 en amplitude, b) performance des filtres mean shift, Frost et FisherMAP, c) image filtrée par fisherMAP et d) ombres extraites pour une $p_{fa} = 2\%$
 e) image filtrée par le mean shift ($h_s=5, h_A=10$) et f) ombres extraites pour une $p_{fa} = 2\%$

Image 2 : cette image est plus grande (360x520), présente une grande zone d'ombre très bruitée par des échos multiples.

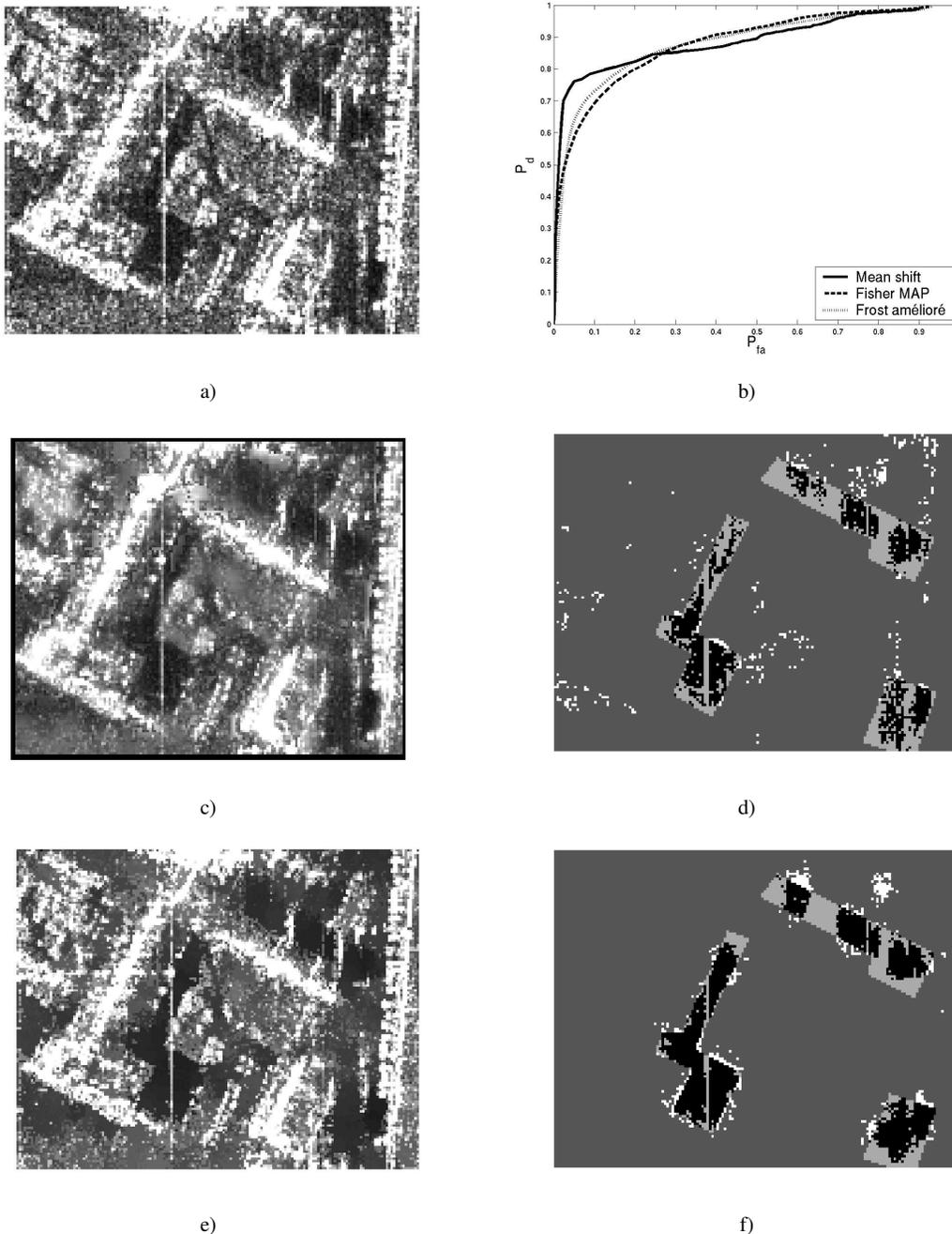


Figure 86 – a) image 1 en amplitude, b) performance des filtres mean shift, Frost et FisherMAP, c) image filtrée par fisherMAP et d) ombres extraites pour une $p_{fa} = 2\%$
 e) image filtrée par le mean shift ($h_s=5, h_A=10$) et f) ombres extraites pour une $p_{fa} = 2\%$

Les filtres paramétriques obtenant les meilleurs résultats sont les filtres Frost amélioré et FisherMAP. Le mean shift, sans faire d'hypothèses sur les statistiques au sein des images donne des courbes COR similaires voire meilleures. En effet, pour une p_{fa} de 2% les ombres extraites sur l'image mean shift sont globalement plus compactes. Ceci provient de la nature du filtre mean shift qui tend à regrouper les pixels similaires. Les fausses alarmes ont également tendance à être regroupées. Dans le cas du mean shift sur l'image 1, beaucoup de

fausses alarmes sont issues des zones frontières où les pixels peu rétrodiffusants de la route ont été associés au mode correspondant à l'ombre.

La p_{fa} de 2% a été choisie parce que les ombres sont correctement détectées pour un taux raisonnable de fausses alarmes. Les p_{fa} où les filtres paramétriques présentent de meilleures détections sont autour de 0.5% pour l'image 1 et 40% pour l'image 2. Sur la figure 87 sont représentées les ombres détectées sur les images filtrées par le meilleur filtre paramétrique (FisherMAP) et par le mean shift pour des p_{fa} avantageuses pour le filtre FisherMAP.

La petite ombre de l'image 1 est détectée sur l'image filtrée par le FisherMAP et ne l'est pas sur l'image filtrée par le mean shift. Pour la deuxième ombre, les fausses alarmes sont situées au niveau du contour gauche de l'ombre où il y a un débordement de 1 pixel. La vérité terrain a été faite manuellement et la délimitation des ombres peut donc présenter des erreurs. Un débordement de 1 pixel correspond, dans ce contexte, à une erreur acceptable. Pour l'image 2, sur les images filtrées seuillées à 4%, le nombre de fausses alarmes est déjà très important. Choisir une p_{fa} de 40% n'est donc pas envisageable.

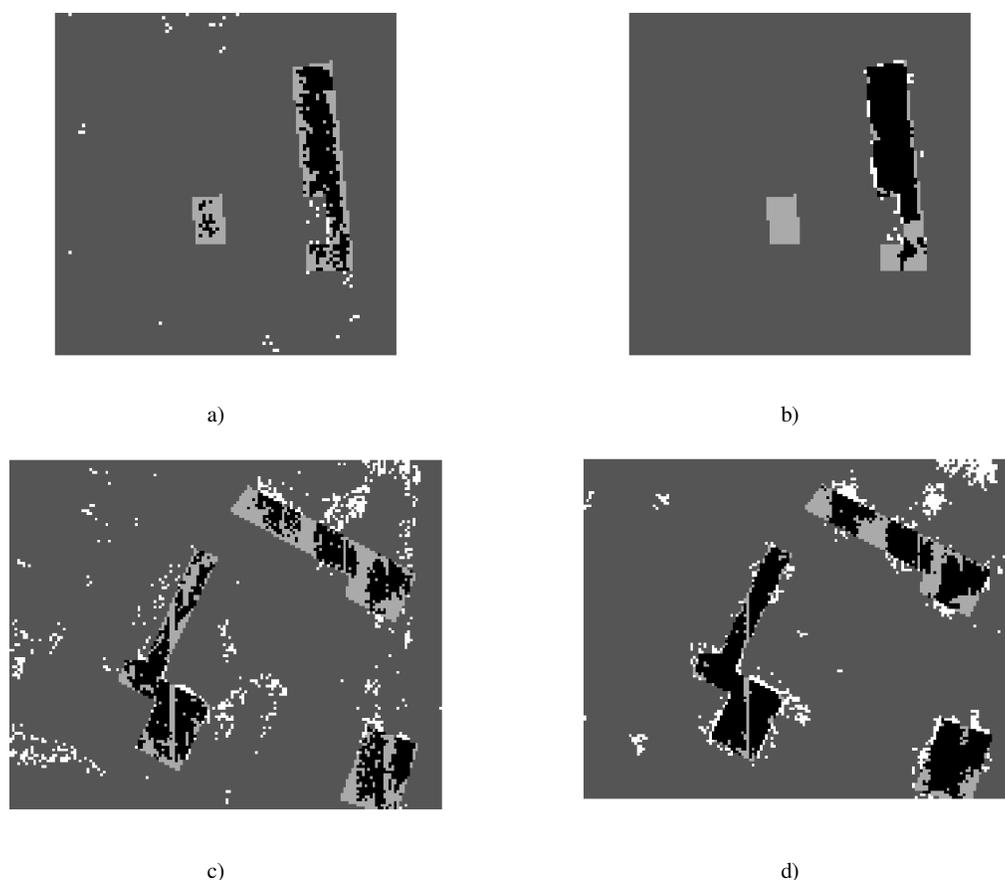


Figure 87 – image 1, $p_{fa}=0.5\%$ ombres détectées sur l'image filtrée par a) FisherMAP et b) mean shift ; image 2, $p_{fa} = 4\%$ ombres détectées sur l'image filtrée par c) FisherMAP et d) mean shift

4.2.7. Conclusion

Le mean shift, filtre utilisé habituellement en imagerie optique, a été introduit en imagerie RSO. Il tente d'associer à chaque pixel le mode local de l'histogramme lissé. Les performances de ce filtre sur images synthétiques avec des distributions Rayleigh-Nakagami sont similaires à celle du meilleur filtre paramétrique. Ces résultats se sont confirmés sur images RAMSES dans le cas de la détection des ombres. Le filtre mean shift se révèle donc

particulièrement robuste car il parvient à filtrer les images RSO sans adaptation. Dans le paragraphe suivant, le filtre mean shift est étendu à l'image de cohérence pour améliorer la détection des ombres.

4.3. Utilisation de la cohérence pour le filtrage des ombres

4.3.1. Intérêt de la cohérence

Les performances d'extraction des ombres sont limitées par la qualité des images : dans le §4.2.6, l'extraction sur l'image 1 se révèle bien meilleure que celle sur l'image 2. L'intérêt du mean shift est de pouvoir travailler sur des données dans des espaces de dimensions élevées (multi images). Il est utilisé également pour la classification de textures à partir de différents paramètres [Georgescu 2003]. Pour améliorer le filtrage nous proposons donc d'introduire la cohérence dans le domaine joint car les ombres y sont également discernables. Les images de cohérence correspondant aux images du §4.2.6 sont présentées sur la figure 88.

La cohérence donne une information sur les similitudes des images complexes du couple interférométrique (cf. §1.3.5). La phase dans une zone d'ombre est du bruit. Sur les deux images d'un couple interférométrique, ces bruits de phase sont indépendants. Par conséquent la cohérence est faible.

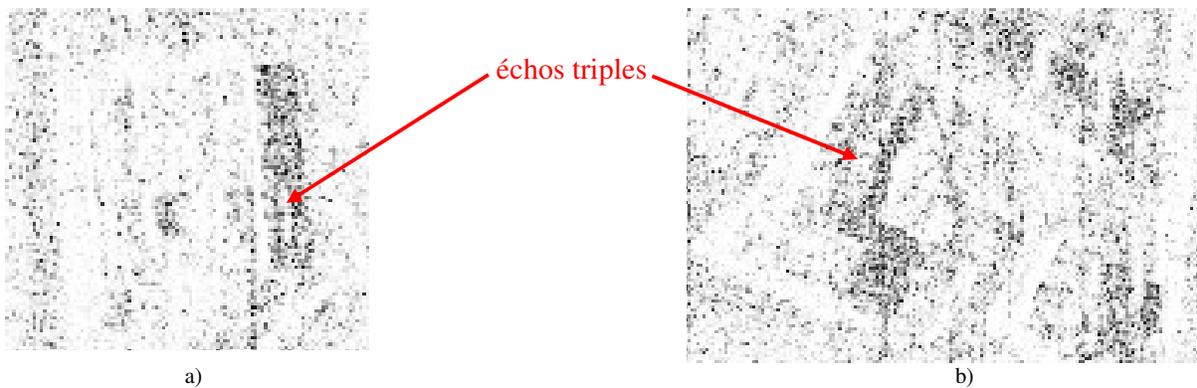


Figure 88 – images de cohérence associées à a) l'image 1, b) l'image 2

Les zones d'ombres sont bruitées par les échos triples. Dans certains cas (figure 88a), ces échos apparaissent nettement dans l'ombre sous une forme linéaire. En général, ils apparaissent plutôt sous forme d'échos ponctuels dispersés qui élèvent la cohérence moyenne de la zone comme dans le cas du bâtiment au centre de la figure 88b.

Les ombres apparaissent beaucoup moins nettes que la théorie nous le laisserait supposer. L'extraction directe de l'ombre sur l'image de cohérence est visiblement délicate. Néanmoins l'image de cohérence fournit une information supplémentaire pour filtrer les ombres.

4.3.2. Extension du domaine joint à l'image de la cohérence

De la même façon que Commaniciu et Meer ont étendu le domaine de filtrage à l'image, nous étendons ce domaine à l'image de cohérence. La densité estimée en x_j^m s'exprime par :

$$\hat{f}(z_j^m) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_A h_C h_S} K_A \left(\frac{a_i^m - a_j^m}{h_A} \right) K_C \left(\frac{\gamma_i^m - \gamma_j^m}{h_C} \right) K_S \left(\frac{s_i^m - s_j^m}{h_S} \right)$$

avec K_C et h_C le noyau et la largeur de bande dans le domaine de cohérence.

Dans les ombres, la cohérence théorique est nulle. L'équation (35), p28, donne :

$$p(\gamma|\Gamma, L) = 2(L-1)\gamma(1-\gamma^2)^{L-2}$$

Pour un nombre de vue $L=4, 9, 16$ et 25 , cette distribution est représentée sur la figure 89. Lorsque le nombre de vues est égal à 4, la distribution de la cohérence empirique dans les zones d'ombre s'étale entre 0 et 1 avec un maximum vers 0,5. La cohérence moyenne est de 0.45. Augmenter le nombre de vue permet de diminuer la cohérence moyenne mais elle n'est jamais nulle. En multivue $L=4$, la largeur de bande dans le domaine de la cohérence est donc choisie à 1. Les pixels sélectionnés pour le filtrage sont ceux qui se situent dans la bande spatiale et la bande amplitude. Comme la bande de cohérence est choisie égale à 1, les mêmes pixels participent au filtrage de l'image d'amplitude lors du filtrage par le mean shift sur amplitude seulement et par le mean shift étendu à la cohérence. L'image d'amplitude filtrée par le filtre étendu à la cohérence est donc la même que celle filtrée par le mean shift sur l'image d'amplitude seulement. En revanche, le filtrage de la cohérence est "guidé" par les largeurs de bandes spatiales et amplitudes. La cohérence est lissée sur les zones d'amplitude homogène.

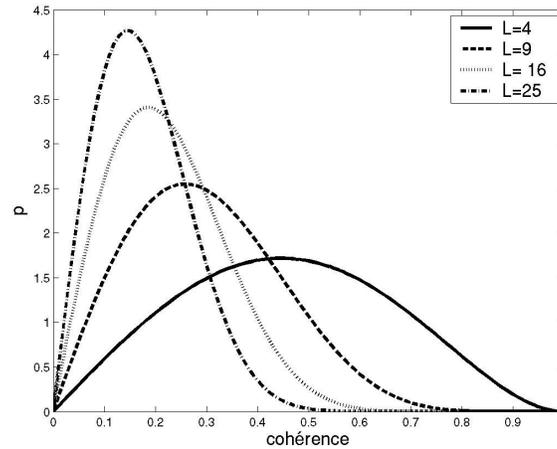


Figure 89 – distribution de la cohérence empirique dans une zone d'ombre pour plusieurs nombres de vues

De la même façon que précédemment, le terme mean shift global est composé de la somme des termes dans chaque domaine. A l'issue du filtrage, deux images (amplitude et cohérence) filtrées sont obtenues.

4.3.3. Fusion des images d'amplitude et de cohérence filtrées

Les ombres peuvent se caractériser par une faible amplitude et une faible cohérence. Les deux images filtrées sont alors fusionnées. En seuillant régulièrement du minimum au maximum l'image de cohérence filtrée et l'image d'amplitude filtrée, deux jeux d'images comportant des ombres hypothétiques sont obtenues. Pour de faibles seuils, les probabilités de fausses alarmes et de détections des ombres seront faibles. Pour des seuils élevés, ces probabilités seront importantes. Les seuils sont notés $S_C = \{s_C\} = \{0, \dots, pas_C, \dots, 1\}$ pour l'image de cohérence et $S_A = \{s_A\} = \{0, \dots, pas_A, \dots, \max(A)\}$ pour l'image d'amplitude. Les images de cohérence et d'amplitude seuillées respectivement à s_C et s_A sont notées $C(s_C)$ et $A(s_A)$.

L'algorithme de fusion consiste à trouver, pour chaque s_C le seuil sur l'image d'amplitude, noté $s_{A,0}$ qui maximise le critère suivant :

$$\Lambda = \frac{C(s_C) \cap A(s_A)}{C(s_C) \cup A(s_A)}$$

C'est-à-dire, pour chaque seuil de cohérence, le pourcentage de pixels qui est étiqueté ombre sur les deux images est maximisé. Pour chaque seuil s_C , l'image fusionnée est notée $F(s_C)$ et consiste en l'intersection de $C(s_C)$ et $A(s_{A,0})$.

L'image fusionnée globale consiste en la somme, pixel à pixel, des images fusionnées aux seuils s_C . Elle est notée F et s'exprime par : $F = \sum_{s_C=0..1} F(s_C)$

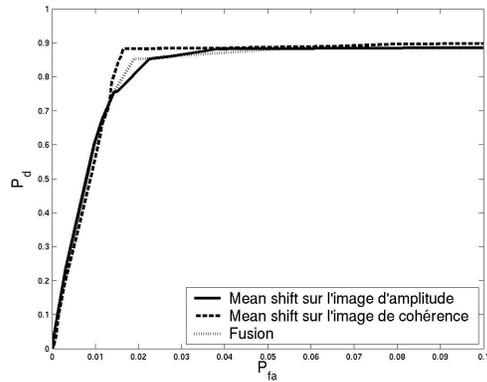
4.3.4. Performance sur images RAMSES

Les performances du mean shift étendu à la cohérence sur images RAMSES sont étudiées sur des images multivues 2x2 (figure 90a et figure 91a). Dans chaque cas, les images de cohérence filtrées et les courbes COR associées sont présentées. Les ombres détectées sur l'image d'amplitude pour une p_{fa} de 2% sont également représentées (figure 90b et figure 91b).

Dans le cas de l'image 1 (cf. figure 90), la détection des ombres sur la cohérence est améliorée. Les fausses alarmes en bas de l'image et en dessous de la grande ombre sont éliminées. En revanche la détection est moins bonne pour l'image 2 (cf. figure 91). Néanmoins, les fausses alarmes sont différentes. Cela confirme bien que l'information au sein des deux images filtrées, même si elle est globalement redondante, n'est pas exactement la même.

Pratiquement, il n'est pas possible de savoir à l'avance si les détections seront meilleures sur l'image d'amplitude ou de cohérence. La fusion des images selon l'algorithme proposé apporte une solution intéressante. En effet, dans le cas de l'image 1, en comparaison des résultats obtenus sur l'image d'amplitude, la fusion permet d'améliorer la détection de la petite ombre et de supprimer la fausse alarme en bas de l'image. Néanmoins, les fausses alarmes en bas de la grande ombre sont toujours présentes. La fusion apporte de meilleurs résultats sur l'image 2. En effet, la détection des ombres est améliorée par rapport à celles obtenues sur les images d'amplitude et de cohérence. De plus une partie des fausses alarmes sur ces deux détections est éliminées. L'algorithme apporte donc un bon compromis et évite de devoir choisir sur quelle image filtrée il faut faire la détection.

image 1 :



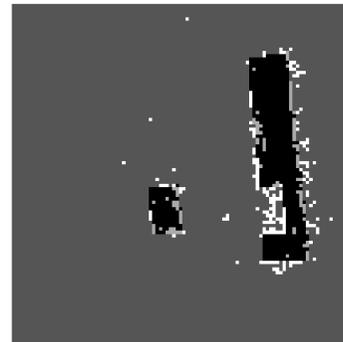
a)



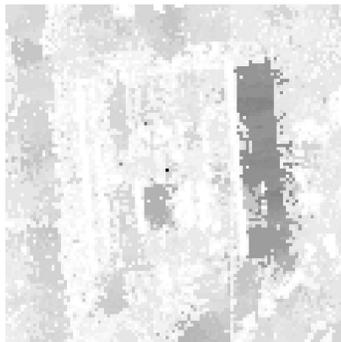
b)



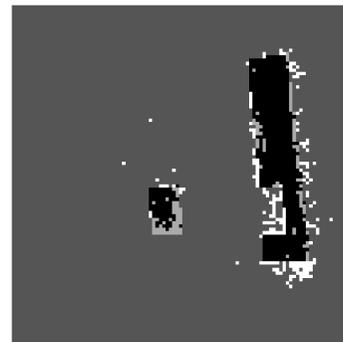
c)



d)



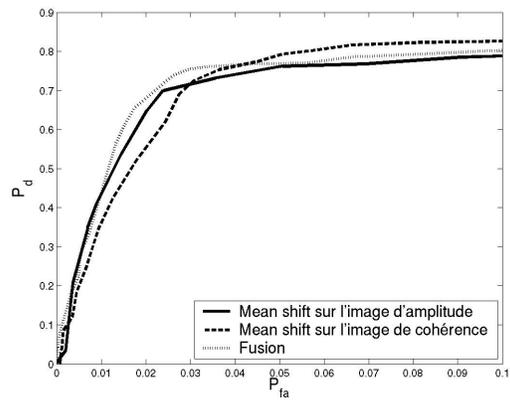
e)



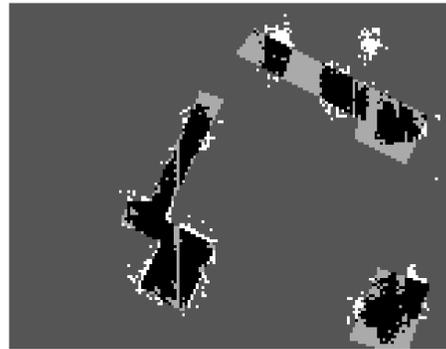
f)

Figure 90 – a) performance sur l'image 1 du filtre mean shift dans le domaine spatial/amplitude/cohérence, b) ombres extraites pour une pfa de 2% après filtrage sur l'image d'amplitude c) image de cohérence filtrée, e) image fusionnée, ombres extraites pour une pfa de 2% après filtrage sur l'image d) de cohérence et f) fusionnée

image 2 :



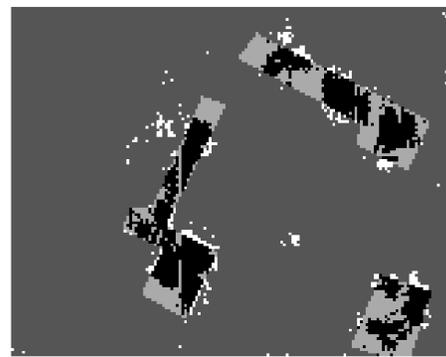
a)



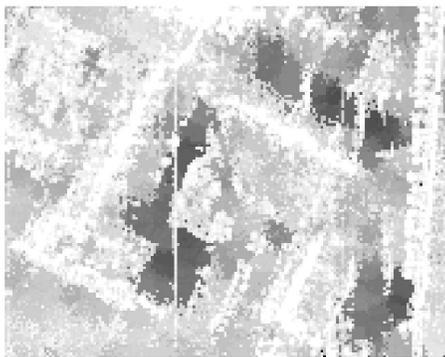
b)



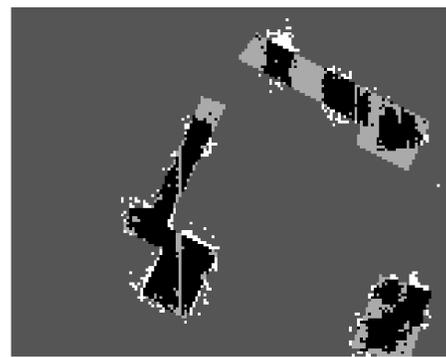
c)



d)



e)



f)

Figure 91 – a) performance sur l'image 1 du filtre mean shift dans le domaine spatial/amplitude/cohérence, b) ombres extraites pour une pfa de 2% après filtrage sur l'image d'amplitude c) image de cohérence filtrée, d) image fusionnée, e) ombres extraites pour une pfa de 2% après filtrage sur l'image e) de cohérence et f) fusionnée

4.3.5. Conclusion

Le filtre mean shift a été étendu à l'image de cohérence pour détecter les ombres. Nous avons vu, grâce aux courbes COR, que l'extraction des ombres sur l'image de cohérence n'est pas forcément meilleure que celle sur l'image d'amplitude mais que l'information disponible sur chaque image est différente. L'algorithme de fusion permet d'obtenir, dans les deux cas présentés, des détections correctes et de supprimer les fausses alarmes. Dans la suite de ce chapitre, les algorithmes d'extraction des ombres, textures et échos forts sont présentés.

4.4. Extraction des signatures de bâtiments en imagerie RSO

4.4.1. Introduction

Après filtrage des images d'amplitude et de cohérence, l'étape suivante consiste à obtenir, à partir des signatures des bâtiments, un jeu de primitives (ombres, textures et échos) utilisables pour la reconstruction. Cette étape est divisée en deux parties. La première consiste à détecter ces primitives ; la seconde est l'extraction et permet, à partir des primitives détectées, d'obtenir des primitives en accord avec les hypothèses choisies sur les bâtiments à reconstruire.

4.4.2. Détection des primitives

Au §2.4, le caractère imprévisible de la présence de textures ou des échos a été mis en évidence. De plus, à cause du repliement, les ombres ne sont pas toujours visibles ou sont bruitées par des échos triples. C'est pourquoi une détection complètement automatique paraît peu robuste et l'intervention de l'utilisateur nous paraît importante pour une bonne détection.

4.4.2.1. Détection des ombres et des textures

Les ombres et les textures n'apparaissent pas à la même amplitude sur toute l'image : le niveau des ombres dépend du rapport signal à clutter local et celui des textures dépend de la longueur d'onde utilisée, des rugosités et de l'orientation du toit.

Le chatolement au sein des ombres et des textures est réduit grâce au filtrage mean shift. Dans le cas des ombres, le filtrage est opéré sur le domaine joint amplitude / cohérence. Les deux images filtrées sont ensuite fusionnées (cf. §4.3). Pour les textures, le filtrage est effectué simplement sur les images d'amplitude avec une largeur de bande plus élevée car le chatolement y est plus important. Sur les images filtrées une croissance de régions est utilisée.

4.4.2.2. Détection des échos forts

La distribution statistique la mieux adaptée aux échos forts au sein d'une zone de rétrodiffusion est la distribution de Fisher (cf. annexe B.4). La contribution de l'écho fort se retrouve en queue de distribution et les valeurs d'amplitudes ont un très grand écart type. Par conséquent, les échos ne peuvent pas être détectés par une croissance de région comme les ombres ou les textures. C'est pourquoi les échos sont détectés par un seuillage sur les images en amplitude.

4.4.3. Description des algorithmes d'extraction

4.4.3.1. Rappels sur la transformée de Hough pour la détection de droites

La transformée de Hough est une technique de reconnaissance des formes utilisée pour détecter des courbes paramétriques telles que des droites, des cercles ou des ellipses dans une image binaire. L'idée de cette transformée, dans le cas de droites, est de faire passer par chaque pixel de l'image toutes les droites possibles.

Sur le schéma de la figure 92a, une image avec un seul pixel non-nul, noté P , est représentée. Deux droites de pentes différentes passant par P sont également tracées. Chaque droite peut être caractérisée de façon unique par une distance ρ et un angle Ψ où :

- ρ représente la distance entre O et le projeté orthogonale de O sur la droite
- Ψ est lié à la pente de la droite.

Ces deux variables correspondent aux axes de l'image dans l'espace de Hough, notée I_H . A la droite D_1 est associé le couple (ρ_1, Ψ_1) . Il y a une infinité de droites passant par un seul point. La transformée de Hough d'une image avec un seul point ressemble à la figure 92b.

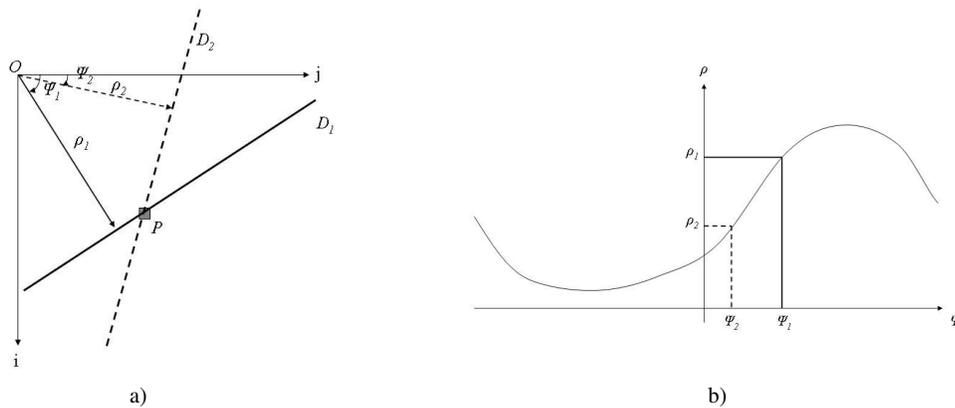


Figure 92 – Principe de la transformée de Hough

Lorsqu'il y a plusieurs pixels non-nuls sur l'image (Figure 93a), une accumulation est effectuée dans l'espace de Hough (Figure 93b). Une structure linéaire se traduit alors par un pic d'accumulation dans l'espace de Hough. Dans l'exemple présenté, ce pic correspond au point M . Détecter ces pics sur l'image permet d'obtenir les droites où apparaissent les structures linéaires à extraire. Le passage de l'espace de Hough à l'espace image est appelé transformée de Hough inverse (Figure 93c).

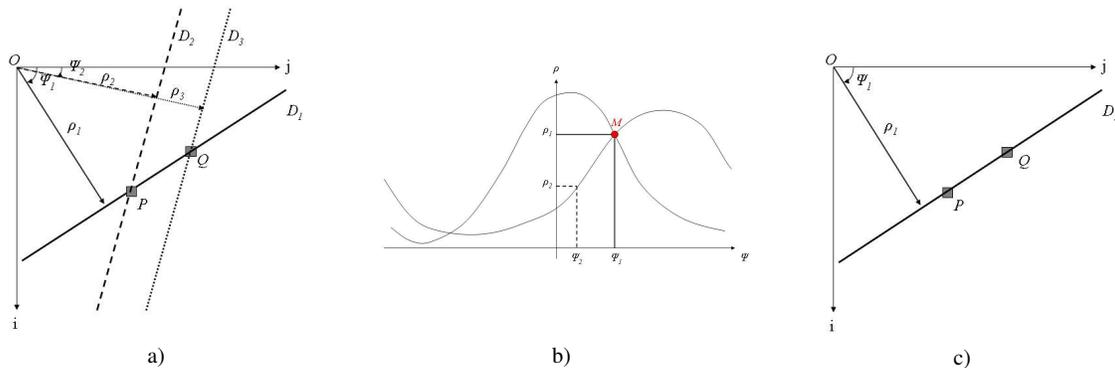


Figure 93 – a) Transformée de Hough sur deux pixels, b) accumulation dans l'espace de Hough, c) transformée de Hough inverse du maximum d'accumulation

4.4.3.2. Détermination des orientations privilégiées par transformée de Hough

Dans l'introduction de la deuxième partie de cette thèse (cf. p92), il est montré que, d'après les hypothèses choisies sur les bâtiments (empreintes à angles droits, toits plats ou en escalier), les ombres, textures et échos suivent les deux orientations privilégiées du bâtiment. Dans le cas des ombres, la direction en distance (dans le cas d'un squint non-nul) est également considérée. Ces orientations constituent une information importante pour l'extraction des primitives et sont obtenues par transformée de Hough. Pour une plus grande robustesse, l'orientation du bâtiment est détectée :

- sur l'image binaire résultant de la somme des images des échos, des contours des ombres et des contours des textures détectés.
- sur les deux orientations associées à l'empreinte du bâtiment.

Ensuite, suivant les deux orientations privilégiées, les maxima d'accumulation sont détectés.

Pour faciliter les représentations, un exemple avec seulement des textures extraites est proposé sur la figure 94a. La détermination des orientations privilégiées s'effectue en deux temps. Tout d'abord l'image de Hough est sommée, pour chaque angle, selon les ρ :

$$I_{H,1}(\Psi) = \sum_{\rho} I_H(\rho, \Psi)$$

A cause de l'éventuelle différence d'échantillonnage selon l'axe des distances et celui des azimuts, les deux orientations, associées au bâtiment, ne sont plus orthogonales mais quasi-orthogonales. L'image de Hough est donc repliée pour que les orientations quasi-orthogonales coïncident :

Pour tout $\Psi \in [0; \pi/2[$, il existe un unique $\Phi \in [\pi/2; \pi[$, tel que Ψ et Φ soient quasi-orthogonaux. Les orientations privilégiées sont obtenues par le couple (Ψ_0, Φ_0) qui satisfait :

$$\arg \max_{(\Psi, \Phi)} (I_{H,1}(\Psi) + I_{H,1}(\Phi))$$

et correspondent aux deux modes de la figure 94b. Pour ces deux orientations privilégiées, et pour chaque primitive considérées séparément, l'histogramme des ρ est lissé par méthode des noyaux. La taille de la bande est choisie en fonction de la taille du plus petit bâtiment à reconstruire. Les maxima locaux caractérisant les structures linéaires sont sélectionnés. La transformée de Hough inverse effectuée suivant les deux ou trois orientations privilégiées forment un maillage de l'image binaire (figure 94c).

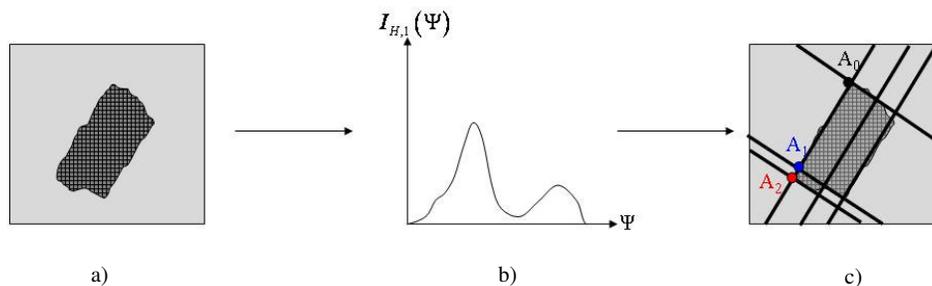


Figure 94 – a) Exemple de texture extraite, b) accumulation selon les ρ pour chaque angle Ψ , c) maillage binaire obtenu par transformée de Hough inverse

4.4.3.3. Extraction des textures et des ombres

L'étape d'extraction vise à pouvoir décrire les primitives extraites par un ensemble de contours fermés, décrits par des segments, qui satisfont les hypothèses effectuées sur les

modèles de bâtiments (bâtiments à angles droits et toits plats ou plats par morceaux). L'algorithme présenté effectue la fusion d'information entre la surface détectée et le maillage binaire obtenu lors de la détection des orientations privilégiées. Une étape de régularisation est ensuite effectuée pour diminuer le bruit sur les extractions.

Soit n le nombre de directions privilégiées que comporte la primitive à extraire. D'après le modèle proposé en p92, dans le cas des textures et des ombres, n vaut respectivement 2 et 3.

Le maillage binaire obtenu à l'issue de la transformée de Hough inverse est constitué d'un ensemble de segments noté S_{Hough} . Ils suivent les orientations privilégiées du bâtiment mais ne correspondent pas tous à la primitive détectée. Cet ensemble de segments est étendu par l'ensemble des combinaisons de segments colinéaires et adjacents. Dans l'exemple de la figure 94c, les segments A_0A_1 , A_1A_2 , A_0A_2 , situés sur la même droite, sont considérés.

Les contours de l'ombre sont également polygonalisés par un algorithme classique de la corde. Cette étape aboutit à un ensemble de segments, noté S_{poly} , qui décrit la primitive détectée mais qui ne satisfait pas les contraintes d'orientations désirées (cf. figure 95b).

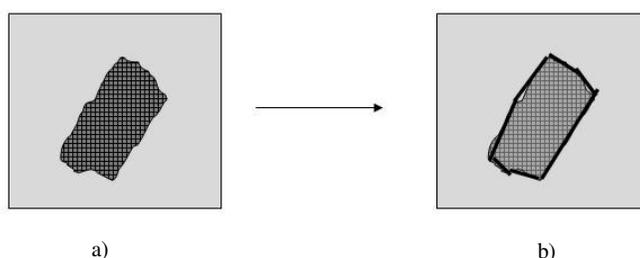


Figure 95 – a) Exemple de texture extraite, b) contours obtenus par polygonalisation

Les propriétés des deux ensembles de segments S_{Hough} et S_{poly} sont complémentaires et ils sont fusionnés. Considérons deux segments s_1 et s_2 . Les extrémités de ces segments sont notées $s_{i,A}$ et $s_{i,B}$, $i=[1,2]$. La distance D entre ces deux segments est définie par :

$$D = \min \left\{ \sqrt{|s_1^A; s_2^A| + |s_1^B; s_2^B|}, \sqrt{|s_1^A; s_2^B| + |s_1^B; s_2^A|} \right\}$$

Pour chaque segment de S_{poly} , seul le segment de S_{Hough} qui minimise la distance précédente est conservé. A l'issue de cette étape, les segments sélectionnés ne forment plus un contour fermé.

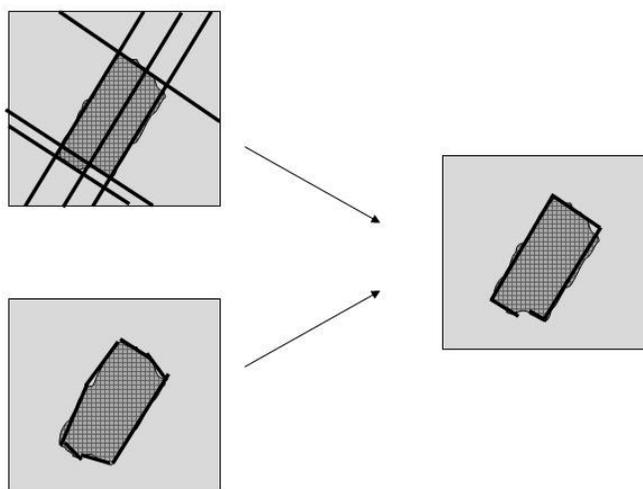


Figure 96 – Fusion des deux ensembles de segments

Pour un bâtiment isolé, le repliement peut tronquer certaines parties des ombres (cf. figure 61, p92). Les extrémités supérieures et inférieures des deux segments qui délimitent l'ombre ne sont pas parallèles. La mesure monoscopique sur ces parties de l'ombre n'est donc plus valide. C'est pourquoi, nous cherchons à éliminer les effets du repliement. Pour cela une régularisation sur l'orientation des contours des surfaces suivant l'axe des distances est effectuée pour imposer des contours parallèles. Le principe de la régularisation est expliqué en détail en annexe E. Les orientations des parties non-fermées de l'ombre ou de la texture extraite sont également obtenues à l'issue de cette régularisation.

Sur chaque zone suivant la même orientation, une régularisation de la largeur suivant l'axe des distances pour les ombres et la deuxième direction privilégiée pour les textures est effectuée.

Les paramètres associés aux régularisations sur l'orientation et la largeur sont notés respectivement Λ_0 et Λ_1 . Plus ces paramètres sont élevés, moins la régularisation attache d'importance aux données et plus la régularisation est forte.

Les ombres et les textures peuvent alors être représentées sous forme d'un ensemble de parallélogrammes. La dernière étape de l'extraction consiste à recalcr ces parallélogrammes sur les primitives détectées en maximisant la densité surfacique de pixels détectés dans chaque parallélogramme.

4.4.3.4. Extraction des échos

L'extraction des échos est effectuée grâce à l'algorithme AROCS (*Algorithme de reconnaissance de formes par optimisation d'un critère surfacique*), appliqué aux bâtiments, proposé dans [Simonetto 2002]. Chaque écho fort est représenté par un rectangle (représenté en bleu sur la figure 97) et est décrit par 5 paramètres, à savoir : la localisation du milieu du segment (i_0, j_0) ; la longueur l ; l'orientation β et l'épaisseur ε . AROCS utilise comme points de départ les L , les T et les croix formés par les échos forts.

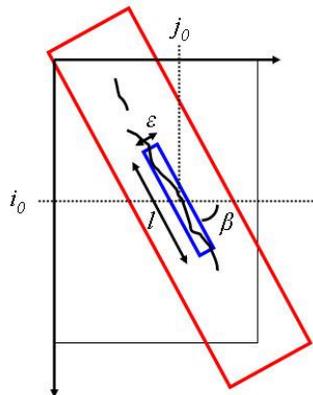


Figure 97 – Principe de la croissance de rectangles

AROCS cherche dans l'espace des paramètres la configuration qui optimise un critère à définir. La recherche des paramètres optimaux est effectuée de façon séquentielle ou par algorithme génétique. La recherche séquentielle fait varier les paramètres pour tester l'ensemble des configurations possibles. Néanmoins, lorsque l'espace des paramètres est grand, cette recherche peut être très coûteuse en temps de calcul. Les algorithmes génétiques permettent de trouver une solution approchée à ce problème d'optimisation en un temps plus raisonnable. Dans le cas des bâtiments, l'auteur montre que la recherche séquentielle fournit

des résultats visuellement meilleurs que la recherche par algorithme génétique car ce dernier nécessite une grande population pour être efficace.

Dans notre approche, les images ne comportent qu'un seul bâtiment. Le nombre d'échos sur les images est donc faible et une optimisation par algorithme génétique n'est pas envisageable. L'optimisation séquentielle est donc choisie. De plus la croissance est effectuée suivant des segments et non des L , T , ou croix car ils ne sont pas visibles lorsque le bâtiment est parallèle à la trajectoire.

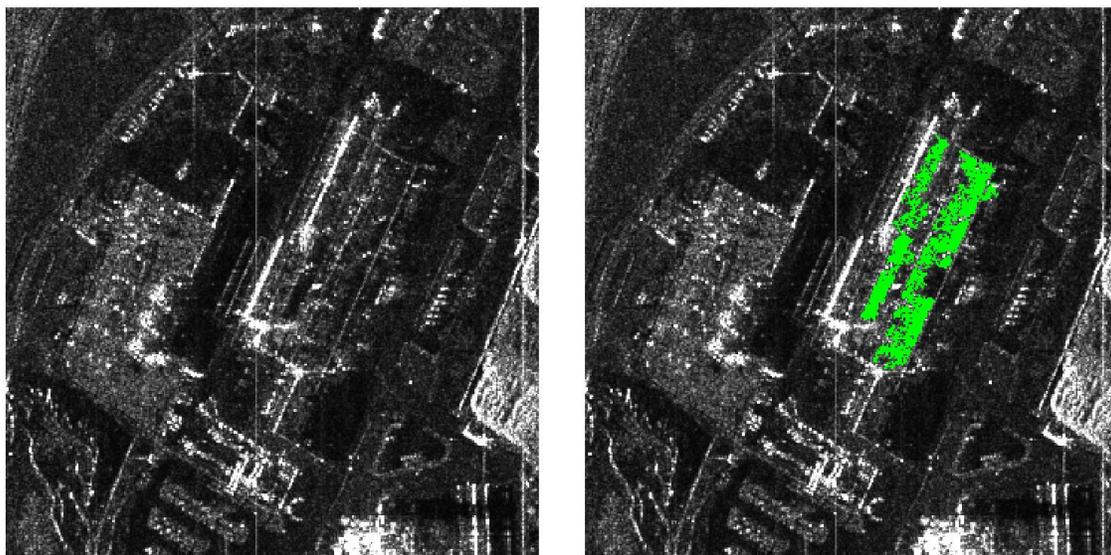
Les points de départ sont les milieux des segments issus du maillage binaire. Le maillage binaire permet également de restreindre l'espace des paramètres car l'orientation β est connue. Le critère à optimiser est défini grâce à deux rectangles (cf. figure 97). Le premier rectangle est représenté en bleu et ses paramètres sont variables. Le second est un rectangle de taille fixe englobant le rectangle variable. Le critère à optimiser est la densité de pixel dans le rectangle variable sur le nombre de pixels dans la couronne entre le rectangle rouge et le rectangle bleu. A l'issue du traitement, si deux segments dans la même direction se recouvrent, celui qui a obtenu le meilleur critère est conservé.

4.4.4. Résultats

Les différents algorithmes d'extraction proposés sont testés sur les images RAMSES. Ces images sont toutes en multivue 2x2 et elles comportent deux types de primitives. Les différentes étapes sont détaillées pour un seul exemple pour chaque primitive. Ensuite, seuls les résultats des extracteurs sont présentés pour diverses valeurs de paramètres.

4.4.4.1. Textures et ombres

Sur la figure 98 sont représentées les différentes étapes de l'extraction de textures dans un cas réel. Les textures détectées (b) ne couvrent pas l'ensemble du toit car le niveau de rétrodiffusion n'est pas homogène.



a)

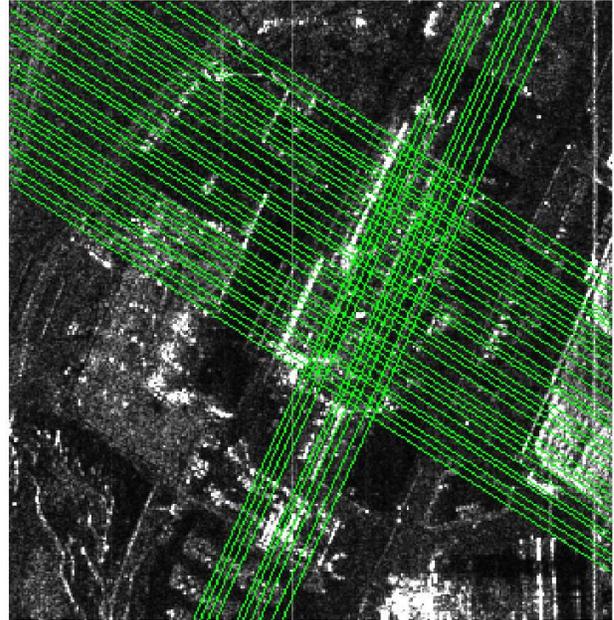
b)

Figure 98 – extraction de textures : a) image d'amplitude
b) croissance de région sur l'image filtrée par le mean shift ($h_S=5$, $h_A=30$)

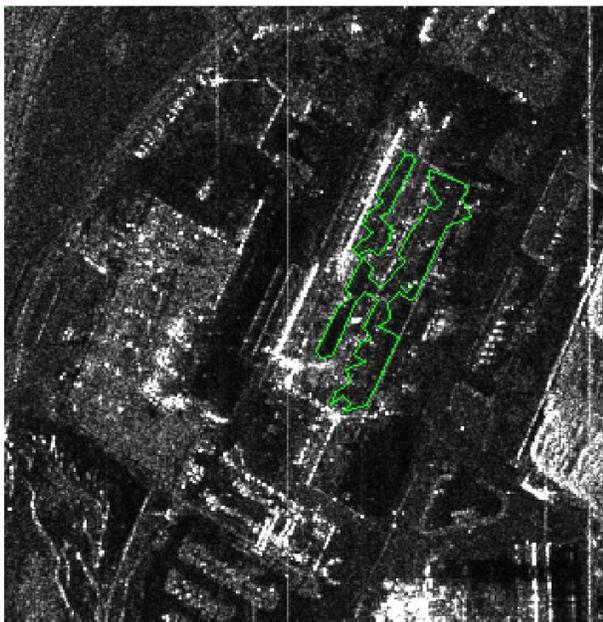
Après fermeture (figure 99a), les hypothétiques contours conformes à la géométrie du bâtiment sont extraits par transformée de Hough (b). Ils forment un maillage de la zone détectée. Les contours de la texture sont polygonalisés (c). Les deux ensembles de segments sont fusionnés (d). Les segments obtenus ne forment visiblement plus un contour fermé.



a)



b)



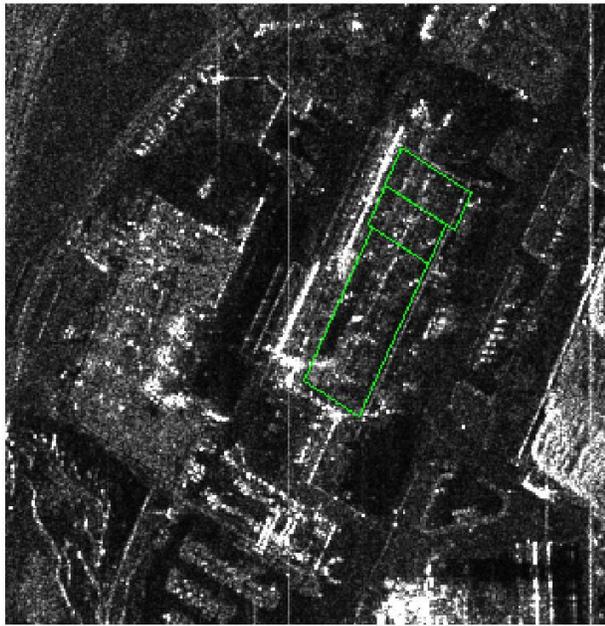
c)



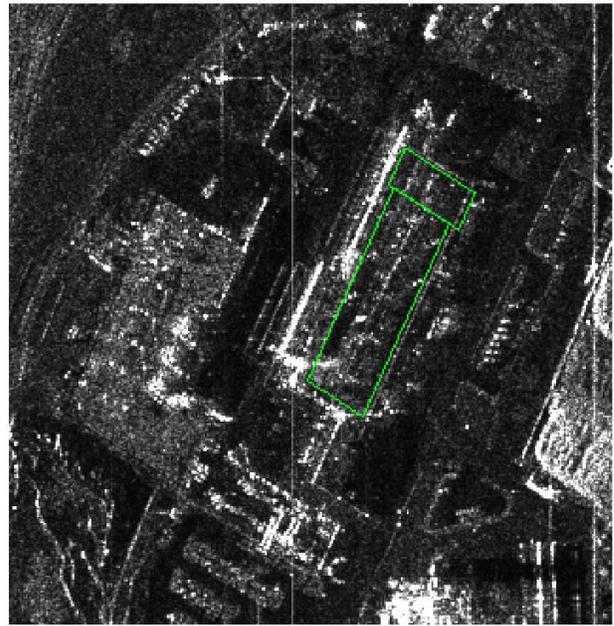
d)

Figure 99 – a) fermeture, b) axes détectés suivant les 2 directions principales du bâtiment
c) polygonalisation de a par algorithme de la corde, d) fusion de b et c

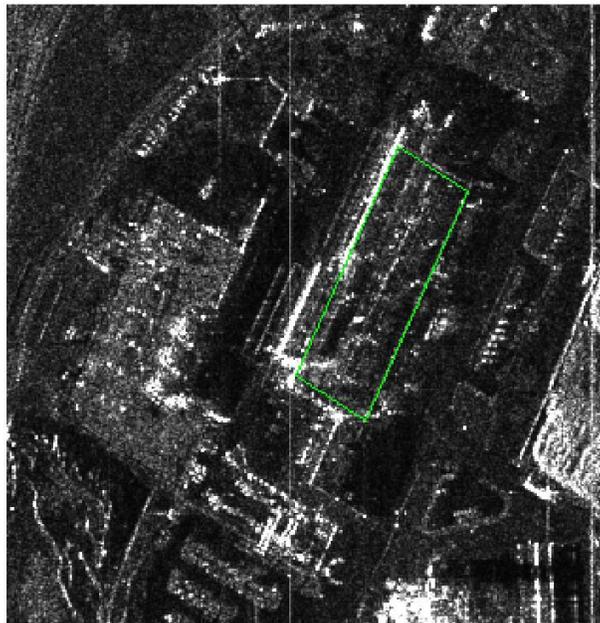
Les résultats de l'extraction sont présentés sur la figure 100 pour trois paramètres de régularisation sur la largeur différents. Pour une faible régularisation, trois parallélogrammes sont détectés (h). Lorsque la régularisation est très forte, il n'y en a plus qu'un (i).



a)



b)



c)

Figure 100 – régularisation avec $\Lambda_0=3$ et a) $\Lambda_1=10$, b) $\Lambda_1=30$ et c) $\Lambda_1=100$

Sur deux autres images (figure 101), où les textures sont bien contrastées, les résultats de l'extraction sont similaires.

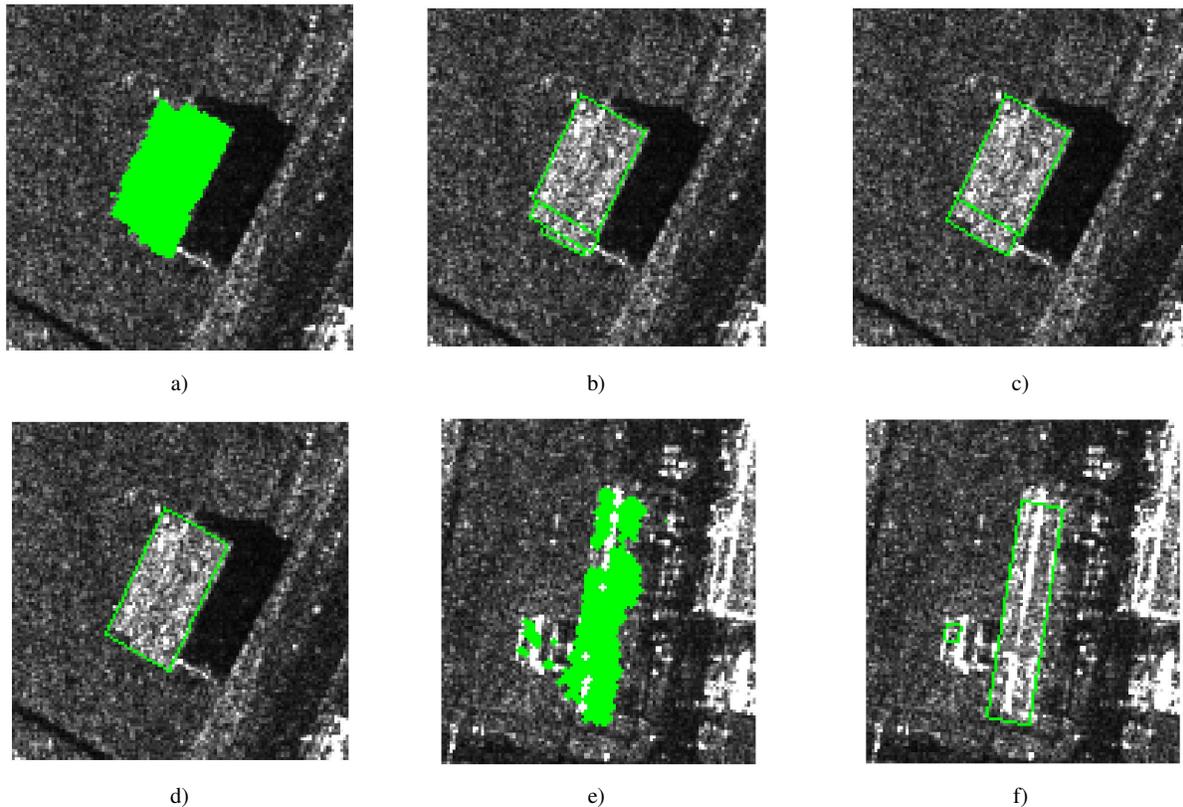


Figure 101 – extraction de textures : a) croissance de région sur l'image filtrée par le mean shift ($h_S=5$, $h_A=40$), textures obtenues après régularisation avec $\Lambda_0=3$ et b) $\Lambda_1=5$, c) $\Lambda_1=10$ et d) $\Lambda_1=15$
 e) croissance de région et fermeture sur l'image filtrée par le mean shift ($h_S=5$, $h_A=30$), f) textures obtenues après régularisation avec $\Lambda_0=3$ et $\Lambda_1=15$

Sur la figure 102, les ombres détectées (a) sont extraites par le même algorithme. Pour deux paramètres de régularisation différents, deux niveaux de détails des ombres sont obtenus. Le niveau de détail le plus fin permet d'effectuer la distinction entre l'ombre et les échos triples

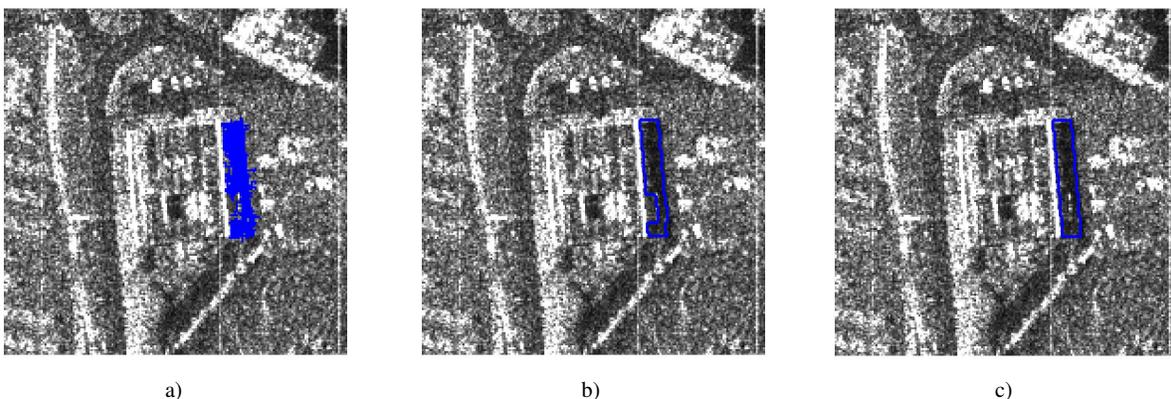


Figure 102 – extraction des ombres : a) croissance de région sur l'image filtrée par le mean shift ($h_S=5$, $h_A=10$), extraction après régularisation avec $\Lambda_0=3$ et b) $\Lambda_1=5$ et c) $\Lambda_1=10$

Sur la figure 103a, la partie inférieure de l'ombre est tronquée par le repliement de la façade. Les segments délimitant l'ombre ne sont plus parallèles. La régularisation sur l'orientation des ombres supprime les effets du repliement (b) pour obtenir une ombre géométriquement adaptée à nos hypothèses.

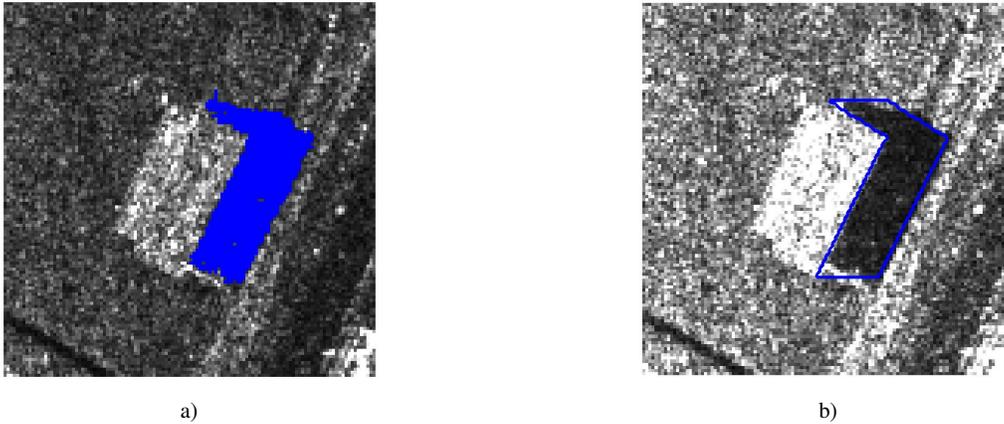


Figure 103 – extraction des ombres : a) croissance de région sur l'image filtrée par le mean shift ($h_s=5$, $h_A=10$), b) extraction après régularisation avec $\Lambda_0=3$ et $\Lambda_1=5$

Les résultats présentés montrent que l'algorithme proposé permet d'extraire correctement les primitives surfaciques si elles sont correctement détectées. Nos tests ont montré que si les détections sont trop mauvaises, l'algorithme d'extraction ne fournit pas de bons résultats. Selon les paramètres de régularisation choisis, les primitives extraites sont plus ou moins fines. L'intérêt majeur de cet algorithme est d'obtenir, pour la reconstruction, des ombres qui sont toujours géométriquement valides.

4.4.4.2. Echos forts

Les étapes de l'algorithme d'extraction des échos forts sont présentées sur la figure 104. L'image d'amplitude (a) est seuillée et les échos sont représentés en rouge (b). Beaucoup d'échos forts environnants sont également détectés. Seuls ceux du bâtiment et de son voisinage immédiat sont conservés. Les positions hypothétiques des échos forts sont ensuite détectées par transformée de Hough de c) puis transformée de hough inverse.

Chaque segment du maillage est utilisé comme point de départ de la croissance de rectangle. Les résultats des extractions pour plusieurs jeux de paramètres sont présentés en e), f) et g). Lorsque l'on joue sur la localisation de l'écho, les extractions sont similaires. Néanmoins lorsque l'épaisseur des échos est choisie trop grande, l'algorithme a tendance à extraire trop d'échos. Certains traversent complètement le bâtiment selon la direction secondaire. Ces échos sont éliminés par un post traitement imposant au moins 50% de pixels détectés sur chaque écho.

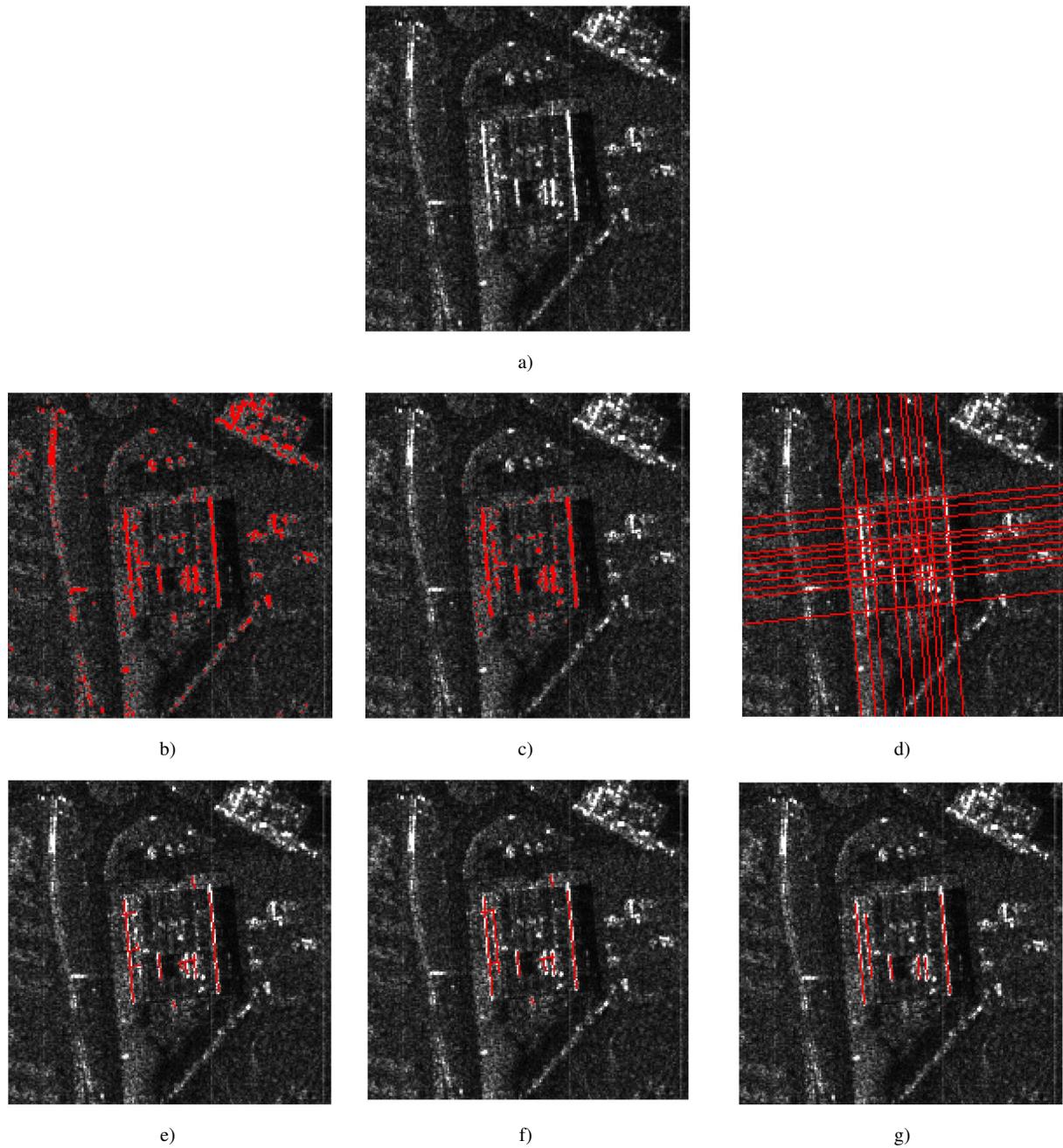
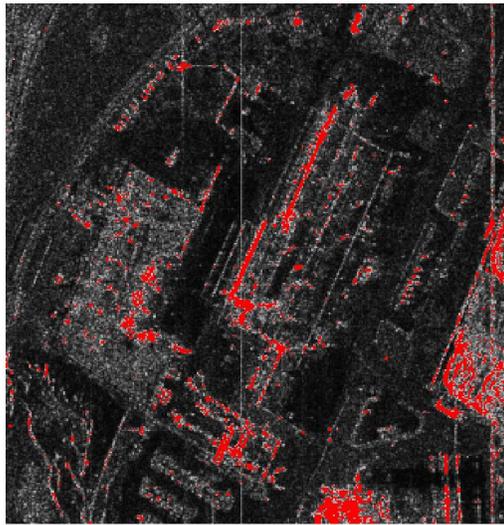


Figure 104 – extraction des échos : a) image d’amplitude, b) seuillage global de l’image d’amplitude, c) sélection des échos du bâtiment, d) transformée de hough inverse de c), résultats de la croissance de rectangle pour des paramètres e) $\beta=1^\circ$, $\varepsilon=1$ pixel, variation de i_0 et j_0 de ± 1 pixel et f) $\beta=1^\circ$, $\varepsilon=1$ pixel, variation de i_0 et j_0 de ± 3 pixels, g) $\beta=1^\circ$, $\varepsilon=3$ pixels, variation de i_0 et j_0 de ± 1 pixel taille mini = 3m

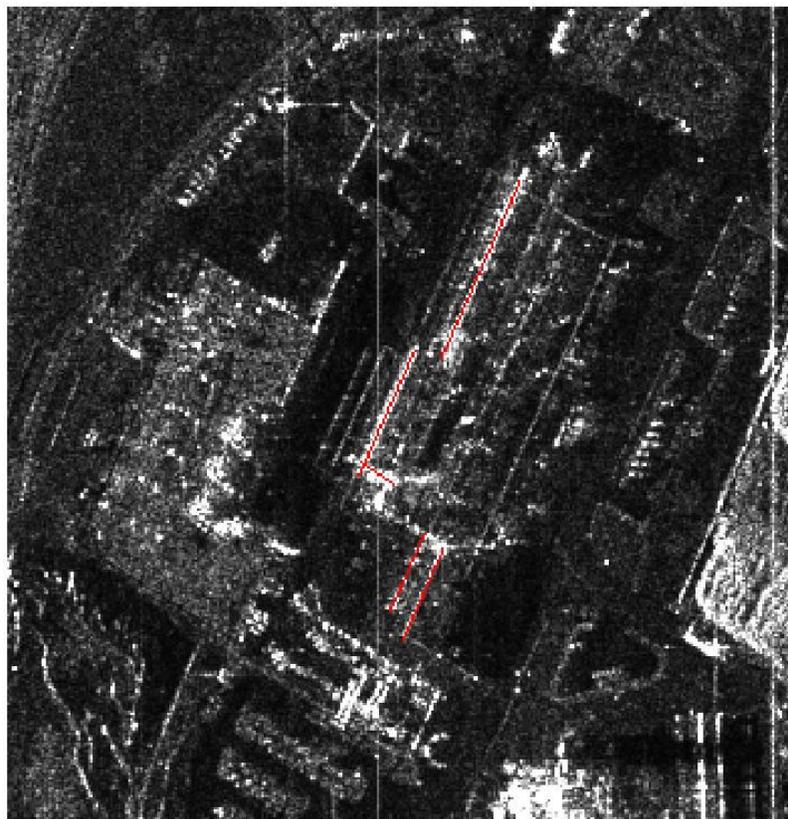
Sur l'exemple suivant, de nombreux échos apparaissent sur ce bâtiment mais seuls les plus brillants sont bien détectés. Les échos du bord du toit ne sont pas détectés (figure 105).



a)



b)



c)

Figure 105 – extraction des échos : a) image d'amplitude avec seuillage global de l'image d'amplitude, b) sélection des échos du bâtiment et c), résultats de la croissance de rectangle pour des paramètres $\beta=1^\circ$, $\varepsilon=1$ pixel, variation de i_0 et j_0 de ± 3 pixels (taille mini = 5m)

Enfin, sur le dernier exemple présenté, les échos doubles sont encore bien extraits.

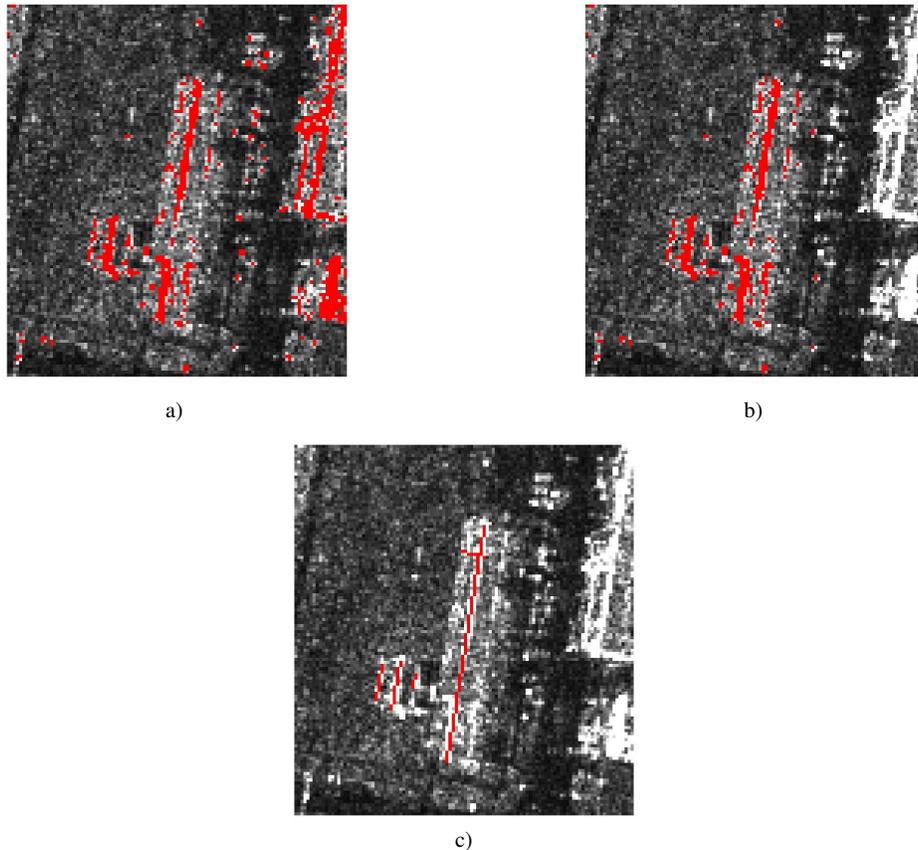


Figure 106 – extraction des échos : a) image d’amplitude avec seuillage global de l’image d’amplitude, b) sélection des échos du bâtiment et c) résultats de la croissance de rectangle pour des paramètres $\beta=1^\circ$, $\varepsilon=1$ pixel, variation de i_0 et j_0 de ± 3 pixels taille mini = 2m

L’algorithme de croissance de rectangles fournit des résultats intéressants car les échos des bâtiments sont bien extraits. Néanmoins le temps de calcul nécessaire pour parcourir l’espace des paramètres est généralement long. Pour réduire ce temps de calcul, l’utilisation de la transformée de Hough pour détecter les orientations principales permet de fixer les angles d’extraction avec une marge de liberté relativement faible. Dans les exemples présentés, elle est de $\pm 1^\circ$. Cette valeur est choisie en rapport avec le pas d’échantillonnage sur les orientations (*tous les degrés*) lors de la transformée de Hough. Les autres paramètres sont relativement similaires d’un exemple à l’autre. L’épaisseur des échos est toujours choisie égale à 1 et la localisation du centre est généralement choisie entre ± 1 et ± 3 pixels.

4.4.5. Conclusion

Les algorithmes d’extraction des signatures de bâtiments en imagerie RSO haute résolution ont été présentés. La première étape consiste à détecter les primitives. Les ombres et les textures sont détectées par croissance de région ou seuillage. Les échos sont détectés par un seuillage. Un algorithme d’extraction des ombres et des textures, basé sur la transformée de Hough, a été proposé. Il permet d’obtenir des primitives géométriquement valides pour la reconstruction. L’algorithme de croissance de rectangles permet d’extraire correctement les échos forts dans les exemples présentés avec des paramètres relativement similaires.

4.5. Conclusion

Dans ce chapitre, la première étape de la chaîne de reconstruction a été détaillée. Tout d'abord les images sont filtrées par le mean shift. Cet algorithme non-paramétrique, initialement dédié aux filtrage d'images optiques, a montré sa robustesse en s'adaptant aux images RADAR. Ses performances obtenues sur images synthétiques ou RAMSES sont similaires voir meilleures que les filtres RSO usuels. Ce filtre a été ensuite étendu à l'image de cohérence pour améliorer la détection des ombres.

A partir des images filtrées, des algorithmes de détection et d'extraction des ombres, textures et échos forts ont été proposés. Nous avons mis en évidence au chapitre 2 le caractère imprévisible des primitives utiles à la reconstruction. C'est pourquoi nous avons choisi une approche semi-automatique pour la détection et l'extraction (seuillage pour les échos forts, paramètres de la régularisation...). L'utilisateur a le choix du type de primitive à extraire et peut évaluer la qualité des primitives extraites en fonction des paramètres des extracteurs. Grâce aux algorithmes proposés, nous avons montré que les paramètres permettant une bonne extraction diffèrent peu et sont relativement intuitifs. Par exemple, lorsqu'une surface (ombre ou texture) est très bruitée, l'utilisation d'une forte régularisation sur la taille de la primitive semble appropriée. Ces algorithmes ont fourni des primitives en accord avec les hypothèses choisies sur les bâtiments à reconstruire.

5. Reconstruction de bâtiments par gestion d'hypothèses en imagerie RSO haute résolution

5.1. Introduction

A ce stade de la méthode de reconstruction, les primitives utiles à la reconstruction sont extraites sous forme de segments (*échos*) ou de contours (ombres et textures) adaptés à la géométrie choisie des modèles de bâtiments, à savoir, des empreintes à angles droits et toits plats ou plats par morceaux. L'ensemble des primitives est supposé incomplet. Nous proposons dans ce chapitre un algorithme de reconstruction de bâtiments sur un unique couple interférométrique par gestion d'hypothèses à partir d'un ensemble incomplet de primitives. Le problème de la reconstruction est alors un problème mal posé. La gestion d'hypothèses, divisée en deux étapes, semble alors adaptée pour combler ce manque d'information. En effet, la première étape consiste tout d'abord à générer l'ensemble des hypothèses à partir des primitives extraites. Ensuite, la validation permet d'éliminer les hypothèses qui ne correspondent pas aux modèles de bâtiments et aux données. Enfin, lorsque plusieurs hypothèses sont obtenues, elles doivent être classées pour aider l'utilisateur à effectuer le choix final.

Dans ce chapitre, l'algorithme de gestion d'hypothèses est tout d'abord présenté. Pour cela une description hiérarchique des bâtiments en imagerie RSO est effectuée. Puis, l'étape de génération d'hypothèses à partir des différentes primitives est détaillée. Différents critères utilisés pour valider ou infirmer les hypothèses générées sont également proposés. Ensuite, une méthode de classement des hypothèses validées est présentée. Enfin, la méthode de reconstruction est validée sur images interférométriques RAMSES haute résolution en bande X.

5.2. La gestion d'hypothèses pour la reconstruction de bâtiments en imagerie RSO

5.2.1. Description hiérarchique d'un bâtiment pour la reconstruction en imagerie RSO

Une description hiérarchique des bâtiments est envisagée selon une approche ascendante. Cette description est représentée schématiquement sur la figure 107. Le plus bas niveau (niveau 0) est représenté par les primitives extraites (échos, ombres et textures).

En utilisant les analyses géométriques et phénoménologiques des §2.2 et 2.4, les primitives extraites peuvent être interprétées et reliées à des éléments surfaciques ou linéiques du bâtiment qui correspondent au niveau 1. Par exemple, une façade peut apparaître sous forme de deux échos ou de textures. Un écho double permet de savoir qu'une arête est formée entre la façade du bâtiment et le sol.

Une partie de bâtiment (niveau 2) consiste en une subdivision du bâtiment, selon l'orientation principale. Les parties sont différenciées selon le nombre et la nature des éléments interprétés qui les composent. Enfin, le niveau 3 correspond au bâtiment complet. Les hypothèses de bâtiment complet sont obtenues par l'ensemble des combinaisons des hypothèses obtenues sur chaque partie.

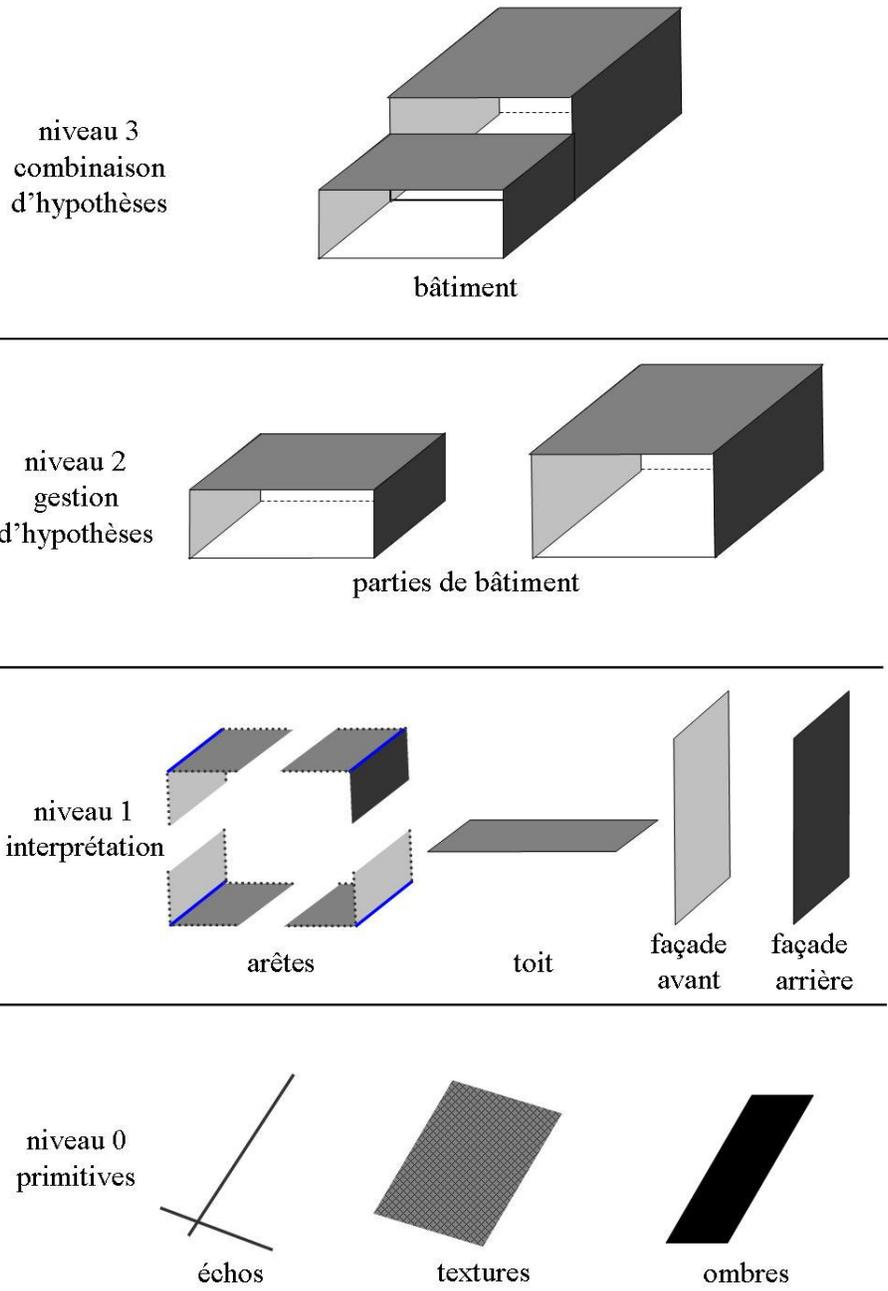


Figure 107 – représentation hiérarchique d'un bâtiment adaptée à l'imagerie RSO

5.2.2. Proposition d'une sémantique pour l'interprétation des signatures

Soit S la sémantique utilisée pour décrire la géométrie des bâtiments à toits plats et plats par morceaux. Les primitives surfaciques sont décrites par deux éléments (début et fin) de cette sémantique.

La sémantique proposée est valable pour décrire les bâtiments à toits plats ou plats par morceaux sur n'importe quel type d'images :

- début toit (DT)
- fin toit (FT)
- début façade avant (DF₁)
- fin façade avant (FF₁)
- début façade arrière (DF₂)
- fin façade arrière (FF₂)
- arêtes (A)
- sol (S)

Le but de l'étape d'interprétation est d'associer les éléments propres aux signatures de bâtiments en imagerie RSO à un élément de la sémantique proposée. Ces éléments propres à l'imagerie RSO sont (les échos triples ne sont pas pris en compte) :

- début ombre (DO)
- fin ombre (FO)
- début texture (DT)
- fin texture (FT)
- écho simple (ES)
- écho double (ED)
- sol (S)

Néanmoins, à une même primitive peut être associé plusieurs éléments du bâtiment. En effet, un écho fort peut correspondre au début et à la fin du toit. Un segment d'ombre peut correspondre non seulement au début ou à la fin d'une façade ombre, mais également à une autre partie du bâtiment lorsque le repliement tronque l'ombre.

C'est pourquoi l'étape d'interprétation n'est pas suffisante pour obtenir une reconstruction directe des bâtiments et des traitements supplémentaires sont nécessaires. L'étape de génération d'hypothèses consiste, entre autre, à gérer ces multiples interprétations.

5.2.3. Représentation d'une partie de bâtiment dans cette sémantique

Chaque partie de bâtiment est tout d'abord représentée sous forme d'une liste chaînée constituée des primitives extraites. Cette liste chaînée débute au niveau du sol derrière le bâtiment pour aller vers le capteur et se terminer au niveau du sol devant le bâtiment.

Pour ordonner les éléments de cette liste, deux géométries sont possibles. Sur la figure 108a, un bâtiment vu en coupe est représenté en pointillés. Seules les primitives visibles sur l'image RSO en amplitude sont ajoutées en tenant compte des éventuelles erreurs sur les mesures. Les ombres sont en noir, les échos doubles en bleu, les échos simples en rouge et le sol en vert.

Sur la figure 108b, les primitives sont ordonnées selon l'axe des distances décroissantes. La liste chaînée est formée des éléments : S – FO – DO – ED – ES – ES – ED – ES – S. Il est visible que, dans les zones de repliement, les jonctions entre les différentes primitives ne correspondent pas du tout à la réalité. Sur la figure 108c, elles sont ordonnées en géométrie

terrain. La liste chaînée est : S – FO – DO – ES – ED – ES – ES – ED – S. Elle correspond beaucoup mieux à la réalité du bâtiment.

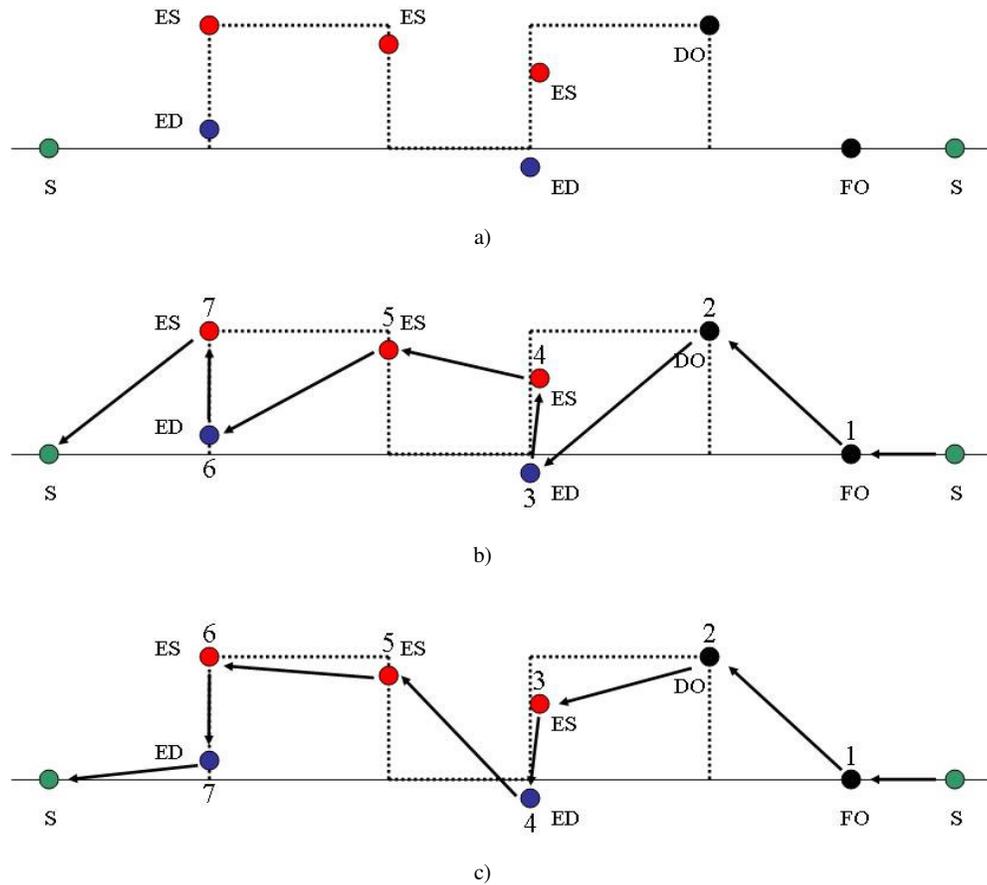


Figure 108 – a) vue en coupe d'un bâtiment et représentation des primitives extraites, ordonnancement des éléments de la liste chaînée des éléments de la signature en géométrie b) RADAR et c) terrain
 code de couleur : vert pour le sol, noir pour l'ombre, rouge pour les échos simples et bleu pour les échos doubles
 code des formes : rond pour primitives extraites

L'exemple de la figure 108c montre également que, même avec une interprétation exacte de la nature des primitives, le bâtiment ne peut pas être décrit totalement : l'ombre entre les primitives, numérotées 4 et 5, n'est pas extraite et le patio du bâtiment ne peut pas être décrit. En plus de gérer les interprétations multiples, l'étape de génération d'hypothèses consistera également à combler les manques d'information et à obtenir un bâtiment conforme aux modèles choisis.

5.2.4. Découpage d'un bâtiment en parties élémentaires

Les bâtiments à reconstruire peuvent être complexes (dépendances, bâtiments en U...). Pour effectuer la reconstruction, il est plus simple de séparer le bâtiment en parties élémentaires. Le découpage du bâtiment peut être effectué en géométrie terrain ou en géométrie RADAR. Nous choisissons de découper le bâtiment en géométrie terrain car cette géométrie est plus représentative du bâtiment.

Pour justifier ce choix, considérons un exemple simple de signature consistant en un toit texturé, un écho double apparaissant sur une portion du bâtiment et un écho simple sur la façade apparaissant dans la même portion. Découper le bâtiment en géométrie RADAR, selon

la direction orthogonale à la direction privilégiée revient à obtenir 5 zones (figure 109a). Lorsque ce bâtiment est découpé en géométrie terrain (figure 109b), seules trois zones sont obtenues. L'écho double et l'écho de la façade sont rassemblés dans la même partie élémentaire. Le découpage en géométrie terrain est donc mieux adapté.

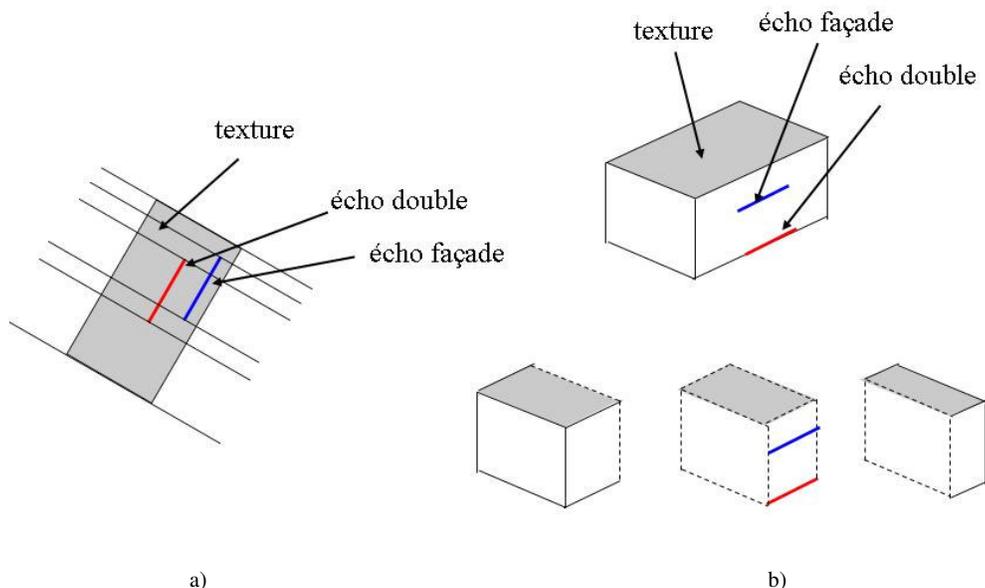


Figure 109 – séparation d'un bâtiment en parties élémentaires en géométrie a) RADAR, b) terrain

Pour les bâtiments composés de plusieurs parties, la description directe sous forme de liste chaînée n'est pas possible. Une liste chaînée doit alors être utilisée pour chaque partie du bâtiment. La génération d'hypothèses est donc effectuée séparément sur chaque partie de bâtiment. Les hypothèses de bâtiments sont obtenues par la combinaison des hypothèses des parties de bâtiments.

5.3. Création de l'arbre d'hypothèses

5.3.1. Définition de l'arbre d'hypothèse initial

L'ensemble de listes chaînées utilisées pour décrire une même partie de bâtiment est appelé arbre d'hypothèses. Chaque liste chaînée est une branche de cet arbre.

Soient s le cardinal de la sémantique S et n le nombre d'éléments de la sémantique utilisés pour décrire le bâtiment. Initialement, aucun *a-priori* n'est utilisé pour générer l'arbre d'hypothèses. Ce qui conduit à avoir s^n descriptions possibles de bâtiments.

Une illustration d'arbre initial est présentée sur la figure 110. En haut est représenté la liste chaînée ordonnée en géométrie terrain et décrite par les primitives extraites. La liste chaînée présentée correspond à une partie de bâtiment dont la signature est composée d'une ombre et d'un écho fort. La lecture de cette liste se fait, comme précédemment, de droite à gauche.

L'arbre initial représenté en dessous part du sol. Sans *a-priori*, le segment fin d'ombre (FO) peut décrire plusieurs éléments du bâtiment. Si l'ombre correspond à une façade arrière, FO correspond au sol. Dans le cas contraire, il peut correspondre au repliement d'une façade d'un autre bâtiment, plus éloigné du capteur et non détectée (DF_1). De même le début de l'ombre (DO) peut correspondre au début d'une façade ombre (et par conséquent la fin d'un toit) mais peut également, à cause du repliement, correspondre à une autre partie du bâtiment

(écho double qui tronque les ombres etc...). L'écho simple, peut correspondre, entre autre, au début du toit ou à un élément de la façade. Pour ne pas surcharger la figure, seule la sémantique associée à la première primitive de la liste chaînée est représentée.

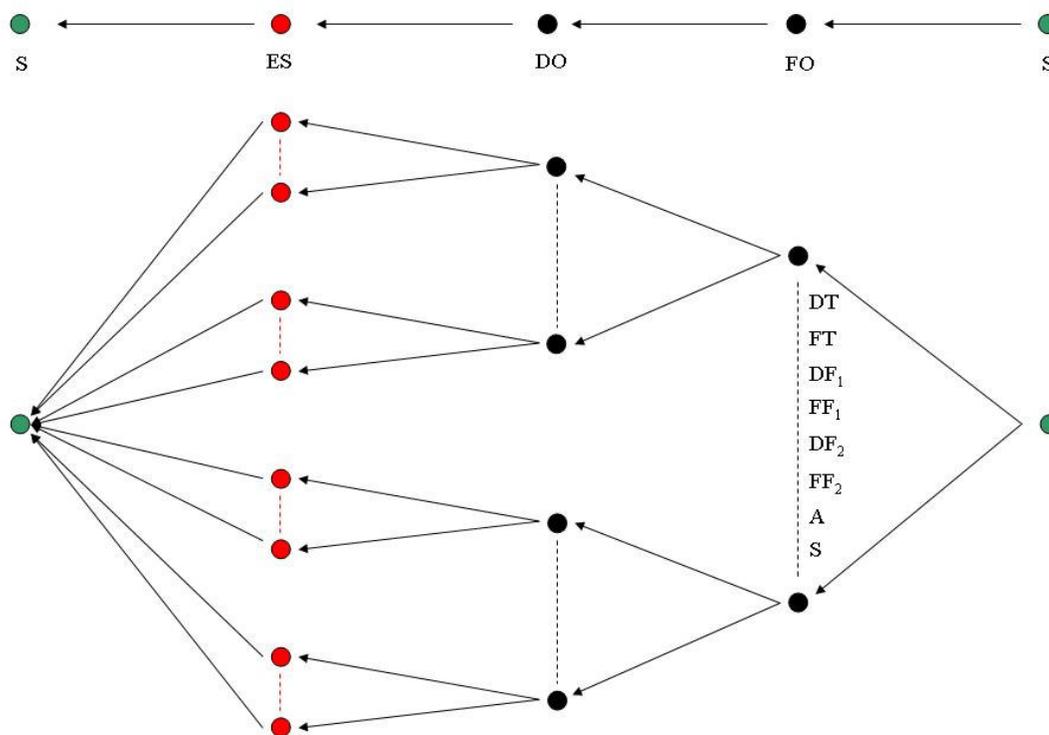


Figure 110 – arbre d’hypothèse initial

5.3.2. Détection de façades avants et arrières

Les façades avants sont détectées à partir des échos et des textures. Les façades arrières sont détectées à partir des ombres. Pour effectuer cette détection, la mesure monoscopique et la différence de hauteur interférométrique entre deux éléments d’une branche sont comparées. Dans le cas de la détection des façades sur les échos et les textures, le modèle présenté au §2.5 est utilisé pour éliminer les effets du repliement et avoir une meilleure estimation de la hauteur interférométrique. Une façade arrière ne peut être détectée que lorsqu’il n’y a pas de repliement qui tronque l’ombre.

La décision est obtenue grâce à un seuillage d’environ 1m50 (de l’ordre de quelques écart-types sur la mesure monoscopique). L’écart-type sur la mesure interférométrique n’est pas considéré pour ce seuillage car il peut être très grand et aboutir à de mauvaises détections. Le seuil choisi est un seuil « dur » et les façades détectées sont considérées comme une information fiable sur les bâtiments à reconstruire car elles sont issues du recoupement d’estimations sur des données différentes. C’est pourquoi, lorsqu’elles sont détectées, elles sont considérées comme des points de passage obligatoires dans l’arbre d’hypothèse. Ce point de passage est unique car la seule façon de relier les deux extrémités d’une façade est par un mur vertical. Dans l’exemple d’arbre proposé à la figure 110, supposons que l’ombre entre les

segments FO et DO soit détectée comme une façade arrière, l'arbre résultant est présenté sur la figure 111.

Un segment, noté FF_2 , correspondant au bas de la façade, est généré. Les éléments générés sont représentés par des carrés. La liaison entre FO et FF_2 est également unique et correspond à une liaison horizontale à la hauteur de FO.

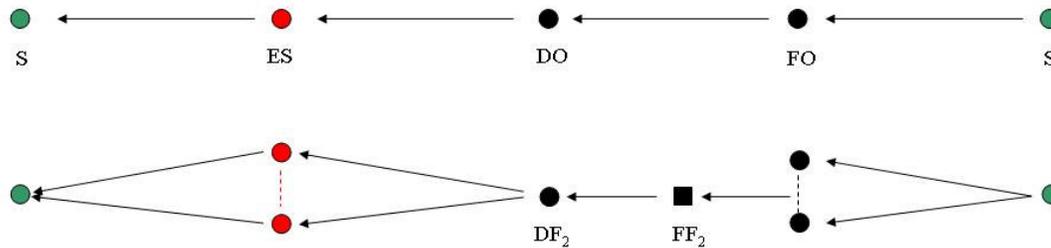


Figure 111 – arbre d'hypothèse après détection des façades

5.4. Génération d'hypothèses de bâtiments

La génération d'hypothèses consiste à générer l'ensemble des listes chaînées compatibles avec les primitives extraites et les modèles de bâtiments pour obtenir des hypothèses sur les parties de bâtiments (niveau 2 de la description des bâtiments). Elle est effectuée localement, entre deux éléments consécutifs de chaque branche, et différemment selon les primitives associées aux éléments considérés. Le cas des échos triples n'est pas pris en compte dans la reconstruction mais est étudié en perspectives au chapitre 6. Lorsque le manque d'information est trop important, la génération d'hypothèses fournit des contraintes géométriques sur la longueur et la hauteur de certains éléments des parties de bâtiments.

5.4.1. Génération d'hypothèses sur les ombres

La génération d'hypothèses sur les ombres est utilisée entre les deux extrémités d'une ombre lorsqu'elle n'a pas été détectée en tant que façade arrière. Dans ce cas, le repliement a tronqué l'ombre et elle est considérée comme non fiable. L'ombre de longueur initiale l (figure 112a) n'est visible que sur une longueur l_0 inférieure à l (figure 112b et figure 112c).

L'origine du repliement ne peut pas être déterminée sauf lorsque le segment FO est au sol (cas de la dernière ombre derrière le bâtiment par exemple) : le repliement a alors obligatoirement lieu sur le début de l'ombre. En effet, si le repliement avait lieu sur la fin de l'ombre, il proviendrait d'un segment plus éloigné et plus haut que le sol (figure 112c). Le repliement peut également avoir lieu simultanément sur les deux extrémités de l'ombre. Des contraintes sur la bâtiment sont alors générées dans la ou les configurations possibles.

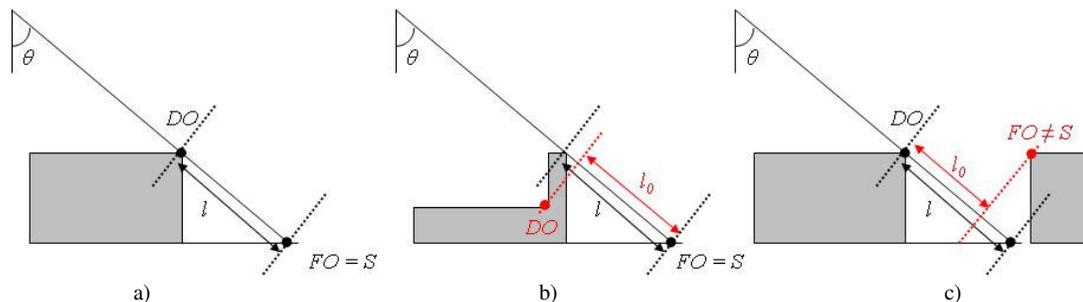


Figure 112 – a) ombre non tronquée par le repliement, b) repliement envisageable avec FO au sol la longueur de l'ombre visible l_0 est inférieure à la longueur de l'ombre réelle l

Sur les schémas qui suivent, les bâtiments, sur une même partie élémentaire, sont supposés être composés de plusieurs parallélépipèdes rectangles numérotés. Les hauteurs des primitives sont calculées par rapport au sol. Les hauteurs des segments DO et FO sont notées respectivement h_{FO} et h_{DO} . Dans le cas où l'ombre est la dernière primitive (la plus éloignée du capteur), FO est supposé au sol ($h_{FO} = 0$). Les schémas sont présentés dans le cas général où $h_{FO} > 0$. On ne peut pas avoir $h_{FO} < 0$ car cela revient à avoir des échos triples sol / mur / sol qui tronquent la fin de l'ombre. Comme les échos triples apparaissent tout le long de la façade, ils bruyeraient l'ombre entière.

repliement sur le début de l'ombre uniquement

L'ombre est tronquée sur son extrémité proche du capteur (DO). Plusieurs cas sont envisageables selon la hauteur par rapport au sol du segment DO (les cas correspondant à des échos triples apparaissant au début de l'ombre, sont envisagés en perspectives au §6.4.1).

- $h_{DO} = 0$

Les échos doubles ont une forte amplitude. Lorsqu'ils tronquent l'ombre (cf. figure 113) deux contraintes sont générées.

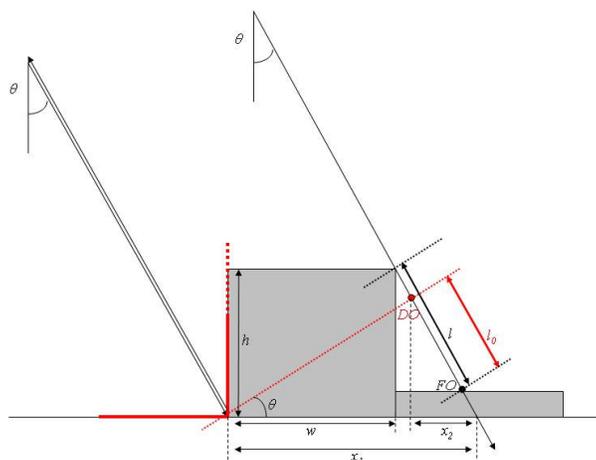


Figure 113 – configuration de repliement sur le début de l'ombre où $h_{DO} = 0$

Les contraintes générées sur h , la hauteur du bâtiment projetant l'ombre, et w , la longueur de ce même bâtiment selon l'axe des distances, sont :

$$h \geq l_0 \cos(\theta) + h_{FO} \quad (80)$$

et

$$w \leq x_1 - x_2$$

x_1 et x_2 s'expriment par :

$$x_1 = \frac{1}{\sin(\theta)} \left(l_0 + \frac{h_{FO}}{\cos(\theta)} \right) \text{ et } x_2 = \sin(\theta) \left(l_0 + \frac{h_{FO}}{\cos(\theta)} \right)$$

D'où :

$$w \leq \left(l_0 + \frac{h_{FO}}{\cos(\theta)} \right) \left(\frac{1}{\sin(\theta)} - \sin(\theta) \right) \quad (81)$$

$$- h_{DO} > 0 \text{ et } h_{DO} < l_0 \cos(\theta)$$

Le bâtiment est complexe et formé de deux parallélépipèdes rectangles notés 1 et 2 (cf. figure 114). Le parallélépipède rectangle 2 projette l'ombre tronquée par l'écho double formé par la jonction entre 1 et 2.

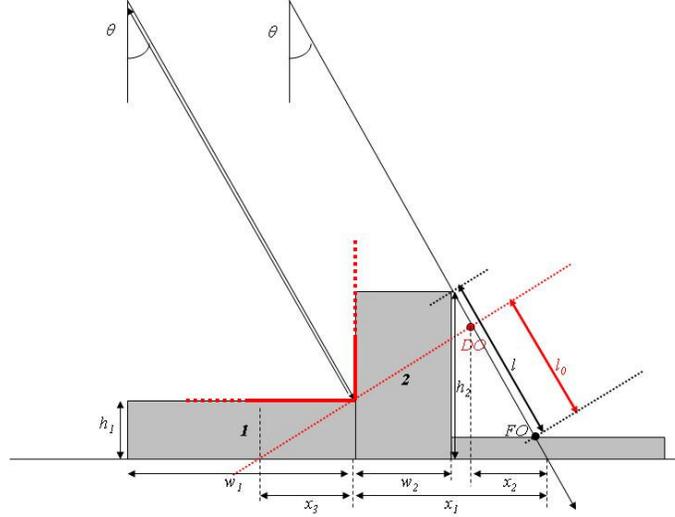


Figure 114 – configuration de repliement sur le début de l'ombre où $h_{DO} > 0$ et $h_{DO} < l_0 \cos(\theta)$

La hauteur du parallélépipède rectangle 1 du bâtiment est connue et vaut $h_1 = h_{DO}$. Les contraintes suivantes sont générées :

$$h_2 \geq l_0 \cos(\theta) + h_{FO} \quad (82)$$

$$w_1 \geq \frac{h_1}{\tan(\theta)} \quad (83)$$

De plus :

$$x_1 + x_3 = \frac{l_0 + \frac{h_{FO}}{\cos(\theta)}}{\sin(\theta)}$$

$$x_2 = l_0 \sin(\theta) + h_{FO} \tan(\theta)$$

D'où

$$w_2 \leq \left(l_0 + \frac{h_{FO}}{\cos(\theta)} \right) \left(\frac{1}{\sin(\theta)} - \sin(\theta) \right) - \frac{h_1}{\tan(\theta)} \quad (84)$$

$$- h_{DO} \geq l_0 \cos(\theta)$$

Ce cas n'est théoriquement pas possible car, pour que la partie proche de l'ombre soit tronquée à cause du repliement, il faut obligatoirement que l'écho soit plus proche et à une hauteur plus faible que $h_{FO} + l_0 \cos(\theta)$. La valeur interférométrique lue au début de l'ombre est alors considérée comme fautive et la façade arrière est générée.

repliement sur la fin de l'ombre uniquement

Ce cas est étudié uniquement lorsque $h_{FO} > 0$. La fin de l'ombre est alors tronquée par un élément plus éloigné du bâtiment (cf. figure 115). La partie élémentaire est constituée de trois parallélépipèdes rectangles numérotés de 1 à 3.

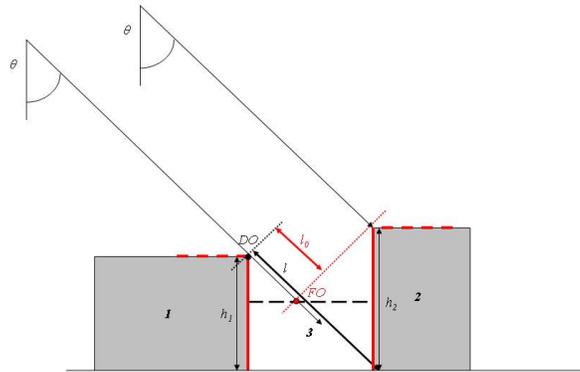


Figure 115 – configuration de repliement sur la fin de l'ombre où $h_{FO} > 0$

Les hauteurs et localisation de la fin de 1 et du début de 2 sont connues grâce aux informations interférométriques.

$$h_1 = h_{DO}$$

$$h_2 = h_{FO}$$

Pour vérifier la fiabilité de cette information sur le bâtiment 1, il faut que la contrainte suivante soit satisfaite :

$$h_1 \geq l_0 \cos(\theta) \quad (85)$$

repliement sur les deux extrémités de l'ombre

La partie de bâtiment est constituée de 4 parallélépipèdes rectangles numérotés de 1 à 4 (cf. figure 116). Ce cas, trop compliqué à gérer n'est pas envisagé. Les segments DO et FO peuvent alors être utilisés, de la même façon que le cas précédent, pour obtenir des informations sur les parallélépipèdes 1 et 3. Aucune information altimétrique fiable sur le parallélépipède rectangle 2 n'est disponible.

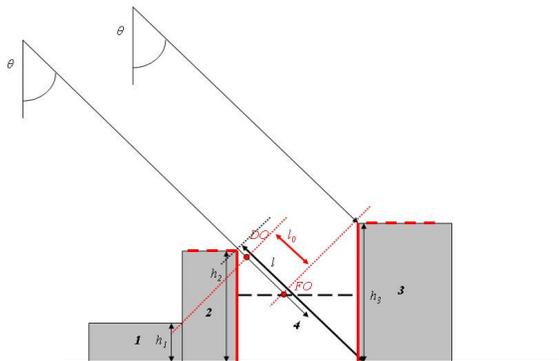


Figure 116 – configuration de repliement sur les deux cotés de l'ombre

5.4.2. Génération d'hypothèses entre deux autres segments

D deux éléments consécutifs, d'une branche de l'arbre, qui ne délimitent pas tous les deux une ombre sont maintenant considérés. Le segment le plus proche du capteur est à la hauteur h_1 . Celui le plus éloigné est à la hauteur h_2 . Un test de confiance sur l'égalité des hauteurs des deux segments est effectué. Les trois cas basés sur la comparaison de h_1 et h_2 sont envisageables. Les hypothèses sont générées pour l'hypothèse de toit plat, plat par morceaux et trou. Pour limiter le nombre d'hypothèses, seuls les cas où il y a une façade avant et 1 ou 2 dénivelés générés sont envisagés.

$h_1 = h_2$

Les segments sont à la même hauteur. Deux hypothèses sont générées (cf. figure 117).

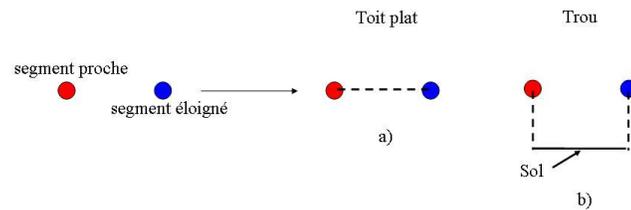


Figure 117 – hypothèses générées si $h_1 = h_2$

$h_1 < h_2$

Le segment proche est plus bas que le segment éloigné. Quatre hypothèses sont générées (cf. figure 118).

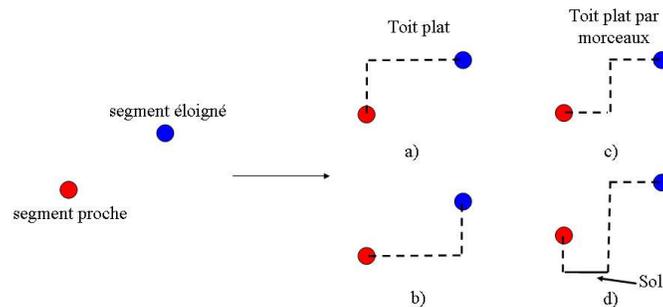


Figure 118 – hypothèses générées si $h_1 < h_2$

$h_1 > h_2$

Le segment éloigné est plus bas que le segment proche. Trois hypothèses sont générées (cf. figure 119).

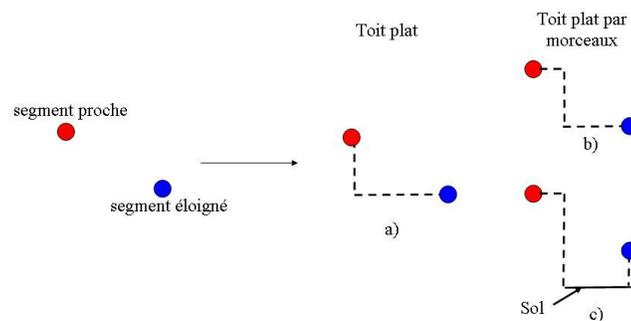


Figure 119 – hypothèses générées si $h_1 > h_2$

A l'issue de l'étape de génération d'hypothèses, un grand nombre d'hypothèses de bâtiments sont obtenues. En reprenant l'exemple de la figure 108, une des branches de l'arbre d'hypothèses de bâtiments obtenues par cette étape est présentée sur la figure 120. De la même façon que précédemment, les éléments générés de la branche sont représentés par des carrés et la lecture se fait de droite à gauche. Dans cet exemple, les façades sont détectées.

Cette hypothèse permet de décrire correctement le bâtiment à reconstruire et elle satisfait les conditions choisies sur les modèles de bâtiments.

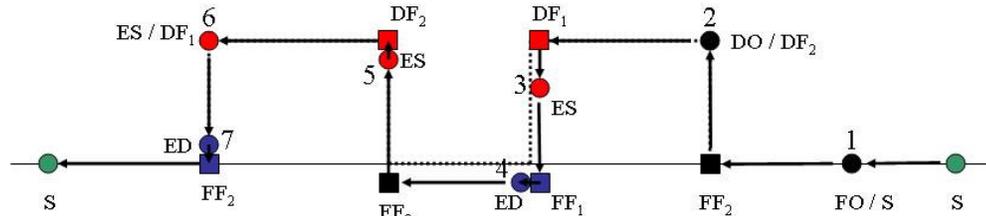


Figure 120 – Exemple d'hypothèse générée
 code de couleur : vert pour le sol, noir pour l'ombre, rouge pour les échos simples et bleu pour les échos doubles
 code des formes : rond pour primitives extraites, carré pour éléments générés

5.5. Validation des hypothèses de bâtiments

La deuxième partie de l'algorithme de gestion d'hypothèses consiste à ne conserver que les hypothèses plausibles. La validation est effectuée sur des critères géométriques et surfaciques (interférométriques).

5.5.1. Gestion des incompatibilités géométriques

Les incompatibilités géométriques proviennent du caractère local de la génération d'hypothèses. Deux hypothèses localement générées sont valides si elles ne présentent pas de recouvrement. Les incompatibilités sont détectées de proche en proche. Par exemple considérons trois échos successifs de hauteurs respectives h_1 , h_2 et h_3 avec $h_3=h_1$ (cf. figure 121). Trois exemples d'hypothèses valides et deux exemples d'hypothèses rejetées sont présentées.

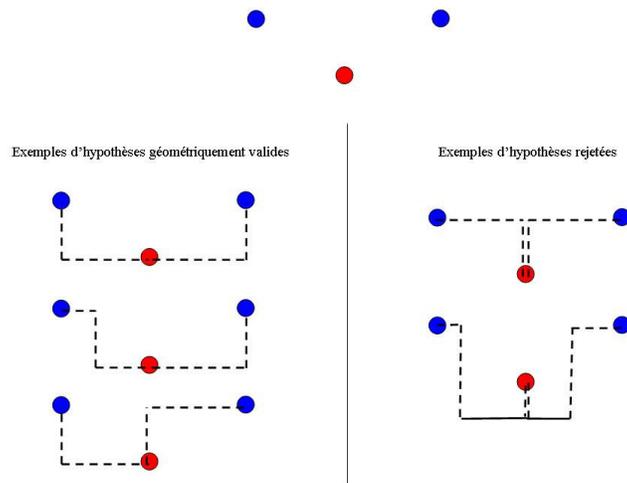


Figure 121 – Exemples de compatibilités et incompatibilités géométriques

Dans le cas des ombres, il ne doit y avoir aucune primitive située entre le début et la fin de l'ombre.

5.5.2. Validation interférométrique

La dernière étape de la validation est basée sur la similarité entre les hypothèses de modèles et les données interférométriques. Tout d'abord le type de toit est déterminé sur chaque élément des branches des arbres d'hypothèses qui ne sont pas des façades ou des ombres tronquées.

Les parties de bâtiments sont découpées selon l'orientation secondaire du bâtiment (orthogonale à l'orientation principale). Si un écho au dessus du sol ou une façade est détecté suivant l'orientation secondaire, sur un ou plusieurs éléments de la branche, alors les sections de bâtiments, correspondant à ces éléments, sont considérées comme des sections à toits plats (cf. figure 122).

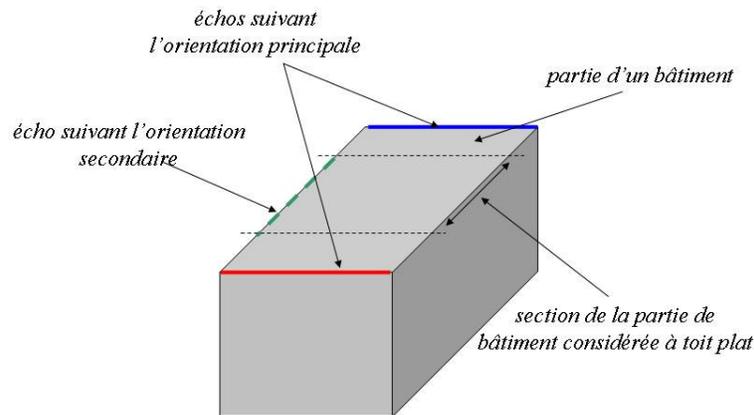


Figure 122 – Utilisation de l'orientation secondaire pour la détermination du type de toit sur une partie de bâtiment

Pour les parties de branches qui ne sont pas des façades et qui ne présentent pas d'information suivant la seconde orientation, les hauteurs des toits sont calculées pour les modèles de toits plats et plats par morceaux en tenant compte de la zone de repliement ou d'ombre générée par cette hypothèse. Sur la figure 123a sont représentés deux échos à deux hauteurs différentes. La surface de toit utilisée sur l'interférogramme pour calculer le modèle à toit plat est représentée en rouge. La zone où intervient le repliement n'est pas considérée. Une erreur quadratique moyenne notée E_{RMS_PLAT} est obtenue.

Néanmoins, dans le cas du toit plat par morceaux la localisation l de la rupture n'est pas connue. C'est pourquoi, toutes les localisations possibles sont testées. Pour chaque position, une erreur quadratique moyenne notée, $E_{RMS_MORCEAUX}(l)$, est calculée. Soit l_0 la localisation qui minimise $E_{RMS_MORCEAUX}$.

L'hypothèse de toit plat est retenue si :

$$E_{RMS_PLAT} < E_{RMS_MORCEAUX}(l_0)$$

Lorsqu'un toit plat est détecté, seules les hypothèses associées au toit plat sont générées. En revanche, si un toit plat par morceaux est détecté, les hypothèses associées aux deux types de toit sont générées car la mesure interférométrique peut être bruitée (toit non-rétrodiffusant, écho triple...) et ne pas correspondre à la réalité du bâtiment.

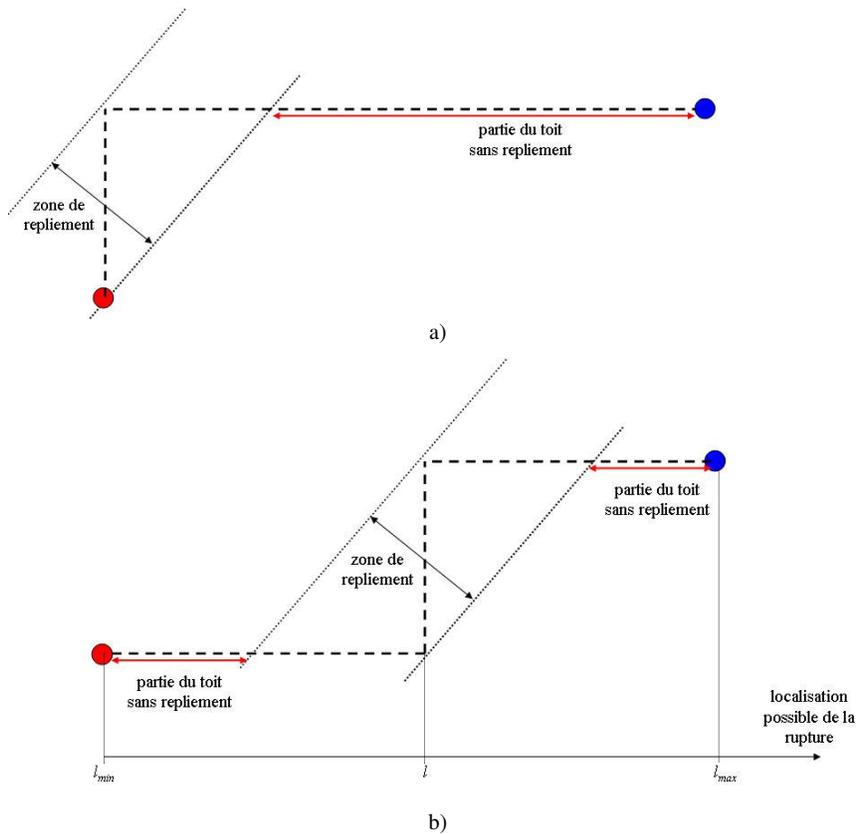


Figure 123 – Zones sur lesquelles sont estimées les hauteurs pour l’hypothèse d’un toit a) plat et b) plat par morceaux

Pour chaque partie plane des toits, la hauteur finale résulte de la fusion des mesures disponibles :

- monoscopiques (sur les ombres et la parallaxe)
- surfaciques (moyenne interférométrique sur le plan)
- échos, façades suivant les deux orientations

Soit n le nombre de mesures disponibles ($n > 1$). Chaque mesure i est associée à une hauteur h_i et un écart-type σ_i . Pour les mesures monoscopiques, σ_i correspond à la précision altimétrique théorique (cf. §2.3.1, p37). Pour les mesures interférométriques, l’écart-type est calculé d’après les hauteurs des pixels constituant la primitive.

A la mesure de hauteur i , on construit la pondération suivante :

$$p_i = \frac{\sum_{\substack{j=1 \\ j \neq i}}^n \sigma_j}{(n-1) \sum_{j=1}^n \sigma_j} \quad (86)$$

La hauteur fusionnée notée h_F , est obtenue par :

$$h_F = \sum_{i=1}^n p_i h_i \quad (87)$$

Cette construction assure que, lorsque σ_i est faible, la pondération associée à la mesure i est importante dans le calcul de la hauteur finale. Connaissant la hauteur fusionnée de chaque partie des hypothèses, les localisations au sol sont affinées.

5.6. Classement des hypothèses

A l'issue de la validation, plusieurs hypothèses peuvent être obtenues. Pour aider l'utilisateur à sélectionner l'hypothèse proposant la « meilleure » reconstruction, un classement des hypothèses est effectué. De part le principe de l'algorithme de gestion d'hypothèses proposé, toutes les hypothèses de modèles passent par les primitives extraites. Le classement doit donc être effectué à partir des éléments du modèle qui ont été générés.

Pour cela une simulation géométrique est effectuée sur les hypothèses obtenues. La simulation géométrique d'une scène consiste à projeter la scène en géométrie RADAR. La simulation géométrique ne tient pas compte des propriétés électromagnétiques de l'onde RADAR.

L'hypothèse présentant la signature simulée la plus proche (géométriquement) des primitives extraites est déterminée. Elle est considérée comme hypothèse référence car elle a le moins d'éléments générés. Les différences entre les signatures des autres modèles et la signature référence sont alors analysées pour déterminer le classement final.

5.6.1. Détermination de l'hypothèse référence

La méthode de détermination de l'hypothèse référence se base sur la minimisation de la complexité stochastique (MCS). Cette approche, introduite par Rissanen [Rissanen 1989], lie la notion de compression au niveau de connaissance que l'on a sur le modèle en partant du principe qu'il est plus facile de compresser quelque chose que l'on connaît bien.

L'idée principale du MCS est, considérant un jeu de donnée Ω , de trouver l'hypothèse H qui comprime au maximum les données. Dans la version basique du MDL, il faut minimiser la somme de deux termes :

$$L(H) + L(\Omega|H) \quad (88)$$

Le premier terme de cette somme correspond à la longueur, en bits, de la description de l'hypothèse. La seconde, est la longueur en bits, de la description des données lorsqu'elles sont encodées avec l'aide de l'hypothèse. Minimiser cette somme revient à trouver un compromis entre la simplicité de l'hypothèse et son adéquation aux données.

Soit p le nombre de segments de l'hypothèse H considérée. La complexité stochastique d'une grille 2D peut s'exprimer par [Galland 2004 ; annexe A. p 138] :

$$L(H) = n \left[\log_2(N) + \log_2(p) \right] + 2 \log_2(\sqrt{p}) + p \left(2 + \log_2(2\bar{m}_x) + \log_2(2\bar{m}_y) \right) \quad (89)$$

où N est le nombre de pixel de l'image

\bar{m}_x, \bar{m}_y sont les tailles moyennes des segments selon les axes de l'image

n est le nombre de sous-graphes eulériens de la grille déterminé par :

$$n = \frac{n_{IMP}}{2} + n_{CI}$$

n_{IMP} est le nombre de nœuds reliés à un nombre impair de segments

n_{CI} est le nombre de graphes isolés ne comportant aucun nœud de multiplicités impaires

Le premier terme de l'équation (89) correspond au nombre de bits nécessaires pour coder les n points de départs et les segments de chaque sous-graphe Eulérien. L'ensemble des sous-graphes décrit la grille et est donc composé de p segments. Les deuxièmes et troisièmes termes de (89) codent respectivement la variance et la moyenne de la taille des segments selon les deux axes de l'image. Les variances des segments sont estimées sur un échantillon de

taille p et peuvent chacune être codée sur $\log_2(\sqrt{p})$ [Rissanen 1989]. La taille du code de la moyenne est obtenue par la logvraisemblance dans [Galland 2004].

Les hypothèses sont obtenues sous forme d'un ensemble de sommets et de facettes et peuvent être représentée sous forme d'un graphe 2D où aucun segments s'intersectent car le bâtiment n'est pas fermé au niveau du sol (il sur le modèle de la figure 124, n'y a pas de facette $A_5A_6A_7A_8$). On peut alors « aplatis » l'hypothèse en représentant à plat et vers l'extérieur du bâtiment les facettes qui correspondent aux murs. Topologiquement l'expression (89) reste valable mais doit être étendue au 3D.

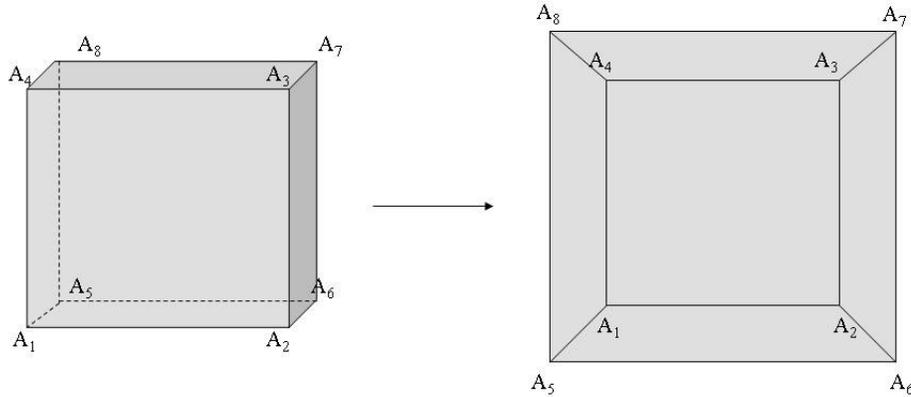


Figure 124 – représentation d'un bâtiment en 3D (à gauche) sous forme de graphe 2D (à droite)

Nous généralisons alors la relation (89) à l'espace 3D en proposant l'expression suivante pour coder l'hypothèse H dans un espace 3D échantillonné :

$$L(H) = n \left[\log_2(V) + \log_2(p) \right] + 3 \log_2(\sqrt{p}) + p \left(3 + \log_2 \left(2 \frac{\bar{l}_x}{pas_x} \right) + \log_2 \left(2 \frac{\bar{l}_y}{pas_y} \right) + \log_2 \left(2 \frac{\bar{l}_z}{pas_z} \right) \right) \quad (90)$$

V est le nombre de voxels de l'espace 3D considéré : la hauteur maximale selon la troisième dimension est définie par la hauteur maximale sur l'ensemble des modèles. pas_x , pas_y et pas_z définissent les pas d'échantillonnage de l'espace 3D selon les trois axes. \bar{l}_x , \bar{l}_y et \bar{l}_z sont les longueurs moyennes des segments, exprimées en mètres, selon les trois directions de l'espace 3D et projetées en géométrie RADAR.

Pour quantifier le terme d'adéquation aux données, il faut comparer les hypothèses de bâtiments aux primitives extraites. Pour cela, un simulateur géométrique est utilisé. Il permet d'obtenir des images sous forme de quatre classes : l'ombre, le sol et les échos et les toits.

Les différences entre la signature simulée de l'hypothèse H et les primitives extraites aboutit à une image composée de surfaces ou segments. Nous voulons comparer ces signatures par type de primitives. Par conséquent, la complexité stochastique de cette image peut être codée par la somme de trois termes (un pour chaque type de primitives) :

$$\begin{aligned} L(H|\Omega) = & n_o \left[\log_2(N) + \log_2(p_o) \right] + \log_2(p_o) + p_o \left(2 + \log_2(2\bar{m}_{x,o}) + \log_2(2\bar{m}_{y,o}) \right) \\ & + n_T \left[\log_2(N) + \log_2(p_T) \right] + \log_2(p_T) + p_T \left(2 + \log_2(2\bar{m}_{x,T}) + \log_2(2\bar{m}_{y,T}) \right) \\ & + n_E \left[\log_2(N) + \log_2(p_E) \right] + \log_2(p_E) + p_E \left(2 + \log_2(2\bar{m}_{x,E}) + \log_2(2\bar{m}_{y,E}) \right) \end{aligned} \quad (91)$$

5.6.2. Analyse des différences de signatures

Le simulateur utilisé ne permet pas d'obtenir une information interférométrique : la hauteur des primitives simulées n'est donc pas disponible. Par conséquent, un classement non optimal est effectué à partir des données interférométriques uniquement. Pour cela, sur les parties différentes des signatures simulées et de la signature simulée référence, la hauteur interférométrique est calculée. Si elle est supérieure au sol, la partie simulée correspondante est considérée comme valide. Plus il y a de parties validées et plus la primitive est probable. Si un simulateur interférométrique était à disposition, l'hypothèse, dont la hauteur de la signature simulée la plus proche des données interférométriques, serait sélectionnée. Sans ce simulateur, si deux hypothèses ont le même nombre de parties validées, nous choisissons l'hypothèse présentant un MDL le plus faible.

5.7. Résultats

La chaîne de reconstruction est testée sur plusieurs zones comportant des bâtiments avec des signatures très différentes. Par exemple, sur la zone 1, les primitives sont la texture du toit et les ombres, alors que sur la zone 2, ce sont la texture du toit et les échos forts. Sur les hypothèses générées, les façades avant sont représentées en blanc et les façades arrières en noir. Le sol est gris clair alors que les toits sont gris foncés.

zone 1 :

Ce bâtiment est constitué de deux parties notées A et B de hauteur 8.45m d'après la BDTopo^{©IGN} (cf. figure 125b). Sur l'image RADAR (cf. figure 125a), ces deux parties sont difficilement distinguables car elles ont toutes les deux les mêmes tailles et niveaux de rétrodiffusion. C'est pourquoi l'extraction des textures sur la figure 125 ne correspond pas exactement au toit. Les ombres sont correctement extraites et représentées sur la figure 125d.

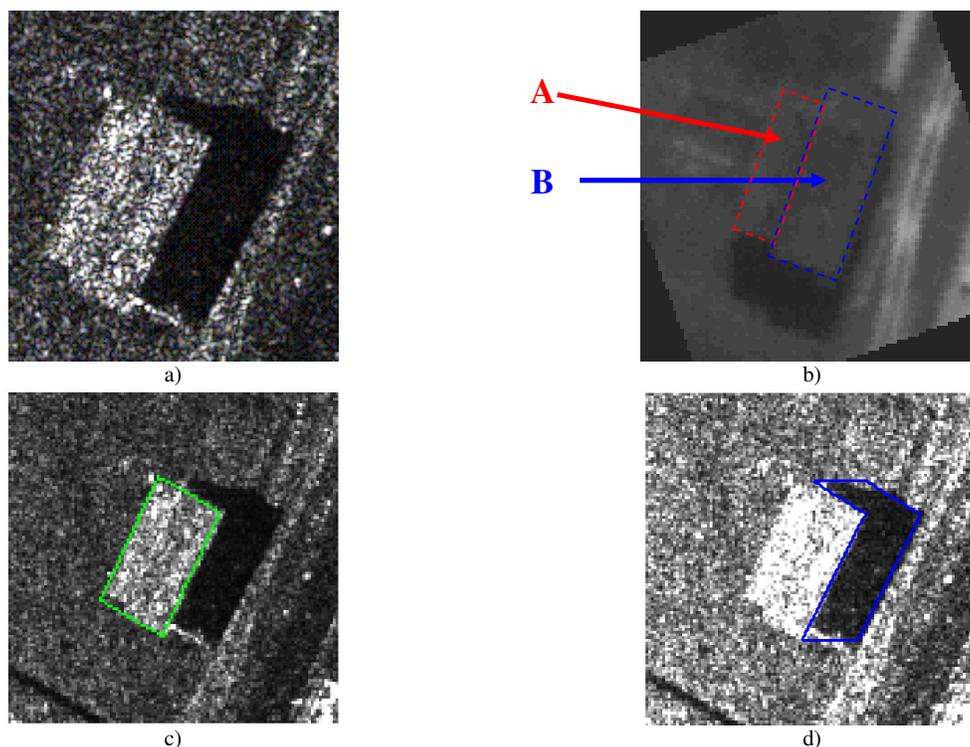


Figure 125 – a) signature RADAR d'un bâtiment, b) image optique correspondante, c) textures et d) ombres extraites

A l'issue de l'algorithme de gestion d'hypothèses, la seule hypothèse obtenue est présentée sur la figure 126. La forme obtenue est cohérente avec la vérité terrain. La hauteur de l'hypothèse obtenue est de 7.77m. Ce qui correspond à une erreur de 68cm.

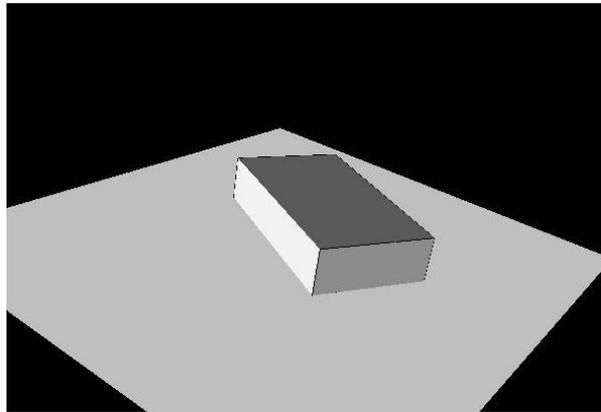


Figure 126 – hypothèse validée par l'algorithme de reconstruction

zone 2 :

Cette zone, dont l'image d'amplitude est présentée en figure 127, contient deux bâtiments proches de forme rectangulaire et à toit plat. L'image optique de la zone correspondante n'était pas disponible. Néanmoins, la BDTopo^{IGN} nous permet d'obtenir une vue schématique de ces bâtiments. Le grand bâtiment (A) a une hauteur de 5.5m et le petit (B) de 3.9m. Sur cet exemple, les ombres sont très bruitées et le coefficient de rétrodiffusion des ombres est proche de celui du bitume. Par conséquent, elles ne sont pas utilisées pour la reconstruction.

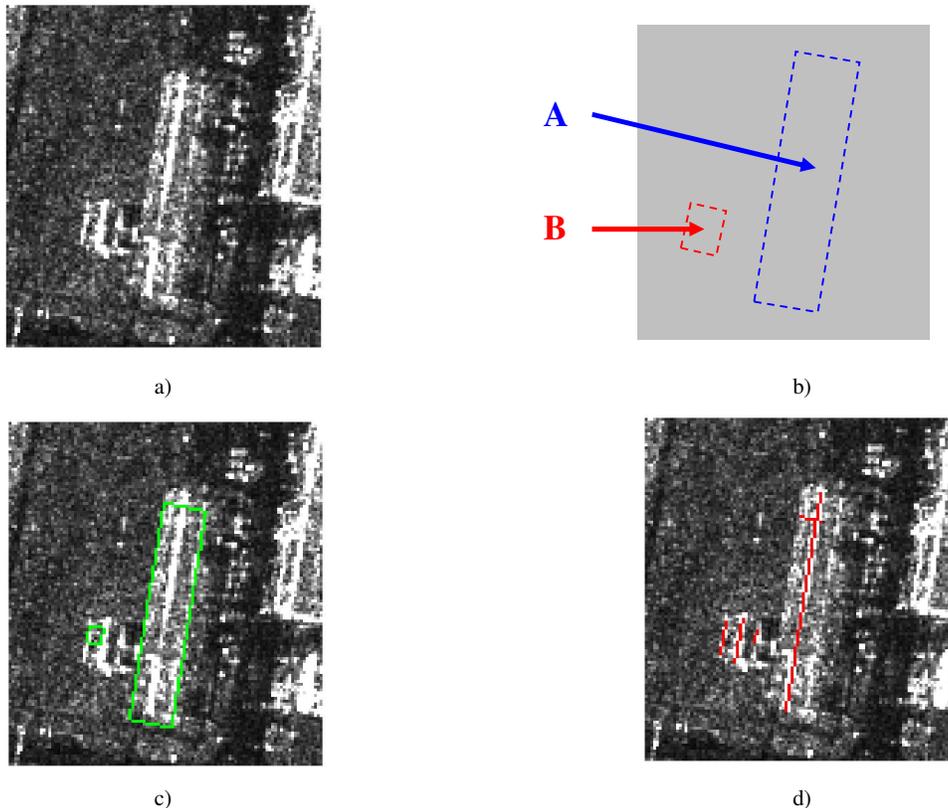
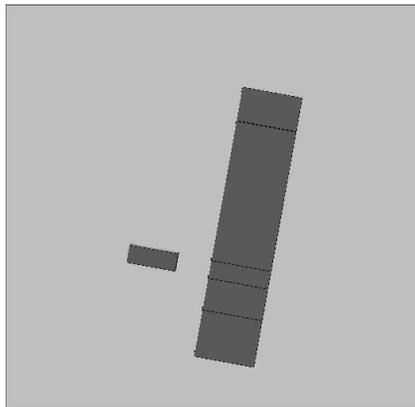


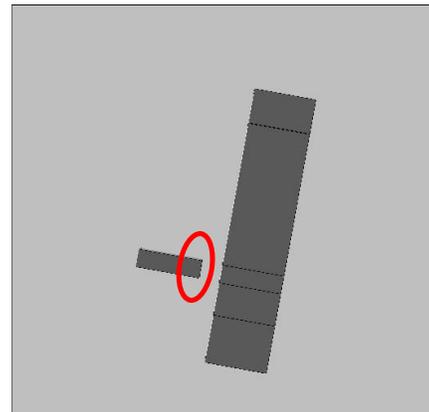
Figure 127 – a) signature RADAR d'un bâtiment, b) vue schématique correspondante, c) textures et d) échos extraits

Deux hypothèses de bâtiments sont obtenues à l'issue de la reconstruction. Sur la figure 128a et b sont représentées ces deux hypothèses avec une vue du dessus. Les lignes noires traversant le bâtiment représentent les parties homogènes de ce bâtiment.

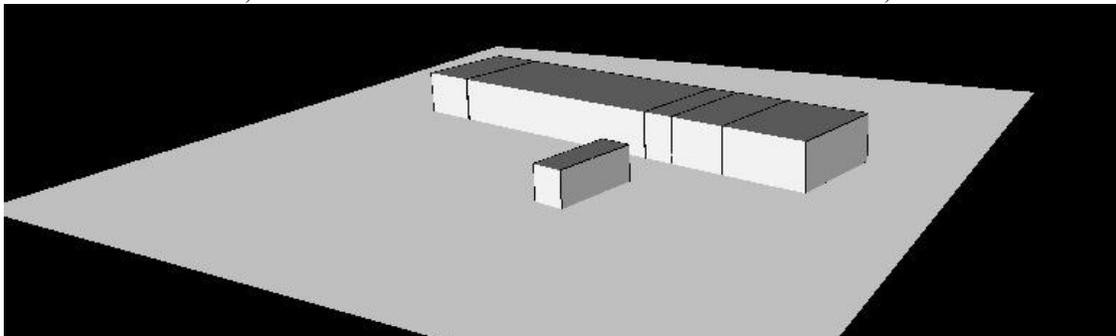
Dans les deux cas le grand bâtiment est bien reconstruit et a la même forme que la vérité terrain. La reconstruction du petit bâtiment est néanmoins imparfaite car les textures n'ont pas été extraites sur le toit complet. Sur la figure 128b, le petit bâtiment reconstruit est légèrement plus long (partie entourée en rouge) que sur la figure 128a. En effet, l'algorithme n'a pas détecté de « toit » plat (le sol) entre les deux bâtiments. Deux hypothèses sont générées en utilisant la figure 119a pour l'hypothèse de toit plat et figure 119b pour l'hypothèse de toit plat par morceaux (dans cette exemple les hypothèses issues de figure 119b et figure 119c sont identiques). Les hypothèses associées au toit plat (figure 128a) et plat par morceaux (figure 128b) sont donc générées.



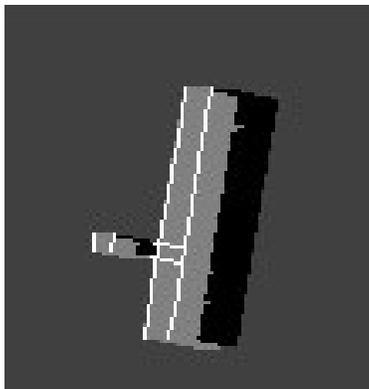
a)



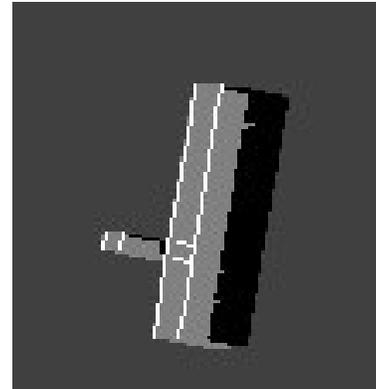
b)



c)



d)



e)

Figure 128 – a), b) hypothèses générées par l'algorithme de reconstruction
 c) vue de face de a)
 d) et e) simulation géométrique de a) et b)

Pour classer ces deux hypothèses suivant l'algorithme proposé au §5.6, le simulateur géométrique est utilisé pour obtenir les deux signatures de la figure 128a-b.

L'hypothèse de la figure 128a est déterminée comme hypothèse référence par le MDL. En effet, les deux grand bâtiments sont identiques et leurs signatures simulées sont donc les mêmes. La différence intervient seulement sur le petit bâtiment. Comme les textures extraites sur ce bâtiment ont une petite surface, l'hypothèse a est privilégiée. La hauteur interférométrique moyenne sur cette zone est au dessus du sol mais cette valeur est peu précise car le moyennage n'est effectué sur 32 pixels seulement. L'hypothèse b est donc classée hypothèse numéro 1. Sur cette hypothèse, les hauteurs des bâtiments A et B sont respectivement 4.8m et 3.60m. Les erreurs respectives sur chaque partie sont de 70cm et 30cm. Sur cet exemple, la reconstruction fournit des résultats géométriquement corrects et avec une faible erreur altimétrique.

zone 3 :

La figure 129a montre la signature RADAR d'un bâtiment dont la signature est beaucoup plus compliquée. La vue optique de la zone est représentée sur la figure 129b. Il s'agit d'un bâtiment (A) de hauteur 9.5m. Sous ce bâtiment, il y a un autre bâtiment (B) de hauteur 12.5m.

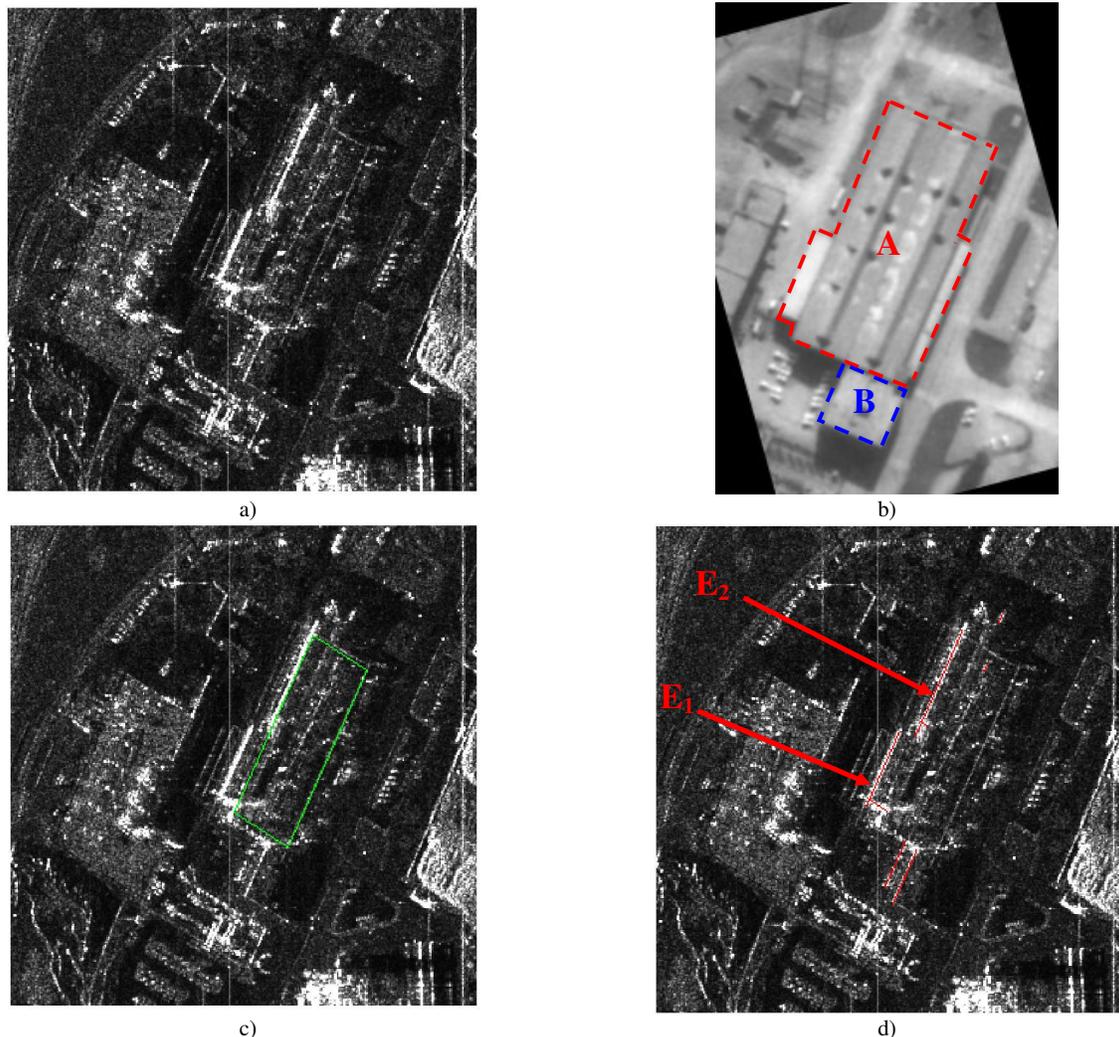


Figure 129 – a) signature RADAR d'un bâtiment, b) image optique correspondante, c) textures et d) échos extraits

Les résultats de la reconstruction sont présentés sur la figure 130. 6 hypothèses sont générées (a-f). Le bâtiment A a été découpé en 3 parties homogènes notées A_1 , A_2 et A_3 . A_1 et A_3 sont associées respectivement aux échos doubles E_1 et E_2 de la figure 129c. La partie intermédiaire A_2 correspond à la zone entre ces deux échos.

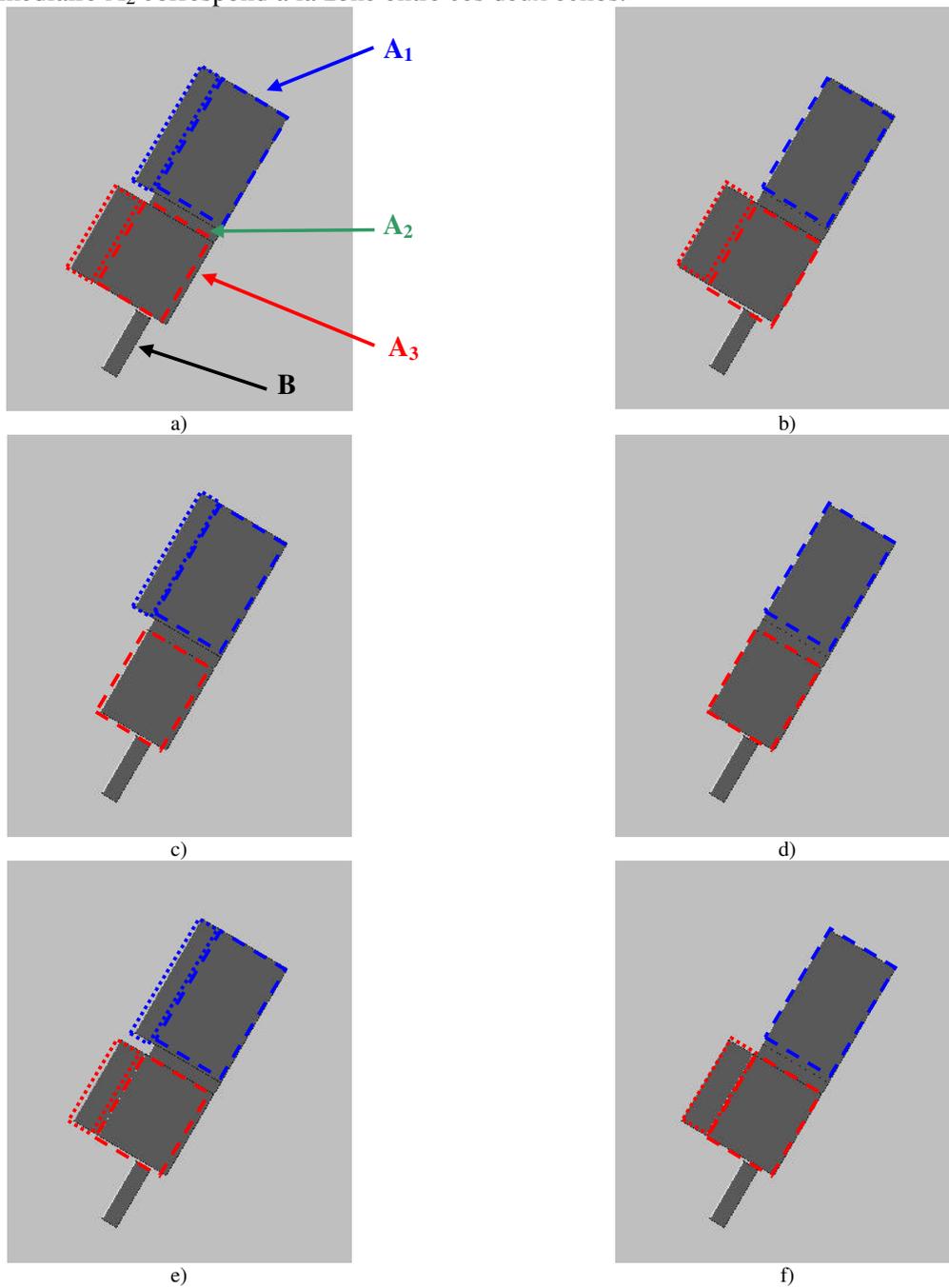


Figure 130 – a-f) hypothèses générées par l’algorithme de reconstruction

Les parties A_1 et A_3 sont elles mêmes découpées en deux parties représentées sur chaque hypothèse par des rectangles de la même couleur mais avec des types de traits différents (traits discontinus ou pointillés). Ces rectangles correspondent aux sections sur lesquels la génération d’hypothèses a été effectuée. Sur les rectangles en traits discontinus, correspondant aux textures extraites, une hypothèse de toit plat est toujours générée. Les sections en

pointillées correspondent à la partie du bâtiment entre le toit et l'écho double. Sur ces sections un toit plat par morceaux est détecté. Cela aboutit dans le cas de la partie A_1 à deux hypothèses différentes et, dans le cas de la partie A_2 à trois hypothèses différentes. Par conséquent 6 hypothèses sont générées. Une vue de face des hypothèses a et c sont représentés sur la figure 131.

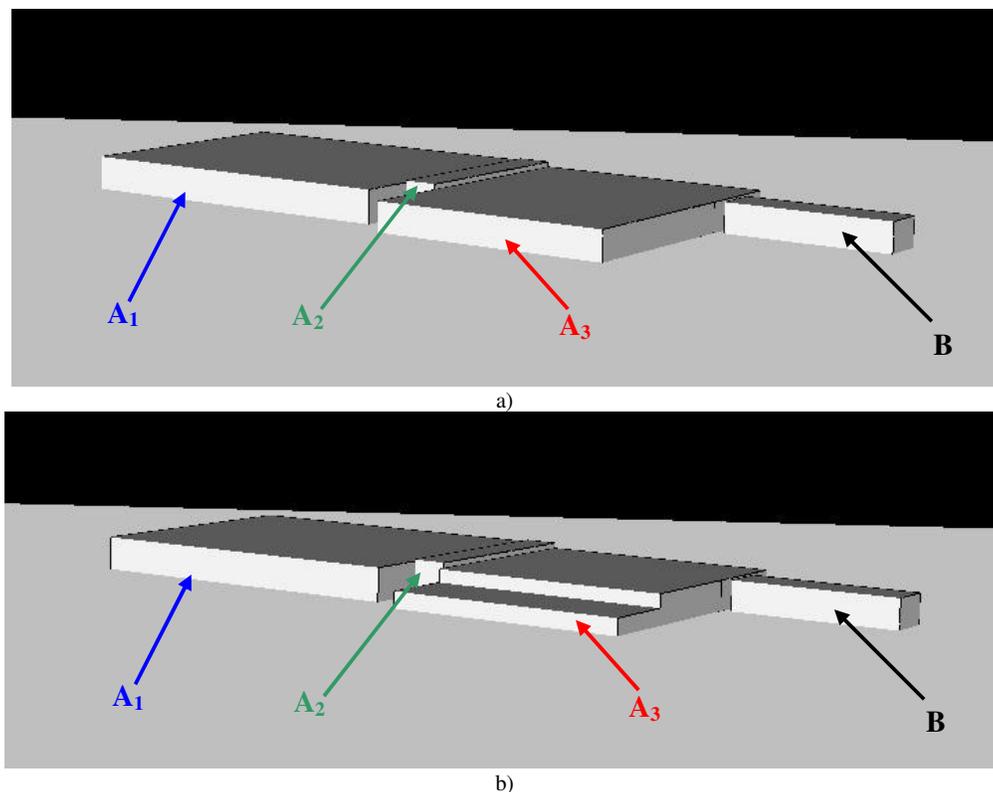


Figure 131 – autres vues de a) l'hypothèse a et b) l'hypothèse c

Pour chaque hypothèse et chaque section, la ou les hauteurs des bâtiments générés sont consignées dans le tableau ci-dessous. Ces hauteurs sont exprimées en mètres.

hypothèse	A_1 (pointillés)	A_1 (discontinu)	A_2	A_3 (pointillés)	A_3 (discontinu)	B
a	5.1	5.1	5.4	4.5	4.5	4.6
b	5.1	5.1	5.4	0	5.4	4.6
c	5.1	5.1	5.4	2.5	4.5	4.6
d	0	5.4	5.4	4.5	4.5	4.6
e	0	5.4	5.4	0	5.4	4.6
f	0	5.4	5.4	2.5	4.5	4.6

L'erreur altimétrique sur cet exemple est de 4-5m. Elle est plus grande que sur les deux cas précédent (au maximum 70cm) car la seule information de hauteur disponible provient de l'interférogramme (il n'y a pas d'ombres ni d'échos du bord du toit extraits).

Néanmoins, géométriquement, l'hypothèse a est conforme à l'image optique : bâtiment à toit plat avec la partie A_3 un peu plus proche du capteur que A_1 . La reconstruction du bâtiment B n'est pas bonne au niveau de la planimétrie et de l'altimétrie. En effet, il n'y a pas assez d'éléments pour délimiter le début et la fin de ce bâtiment.

Les images issues de la simulation pour le classement des hypothèses sont présentées sur la figure 132.

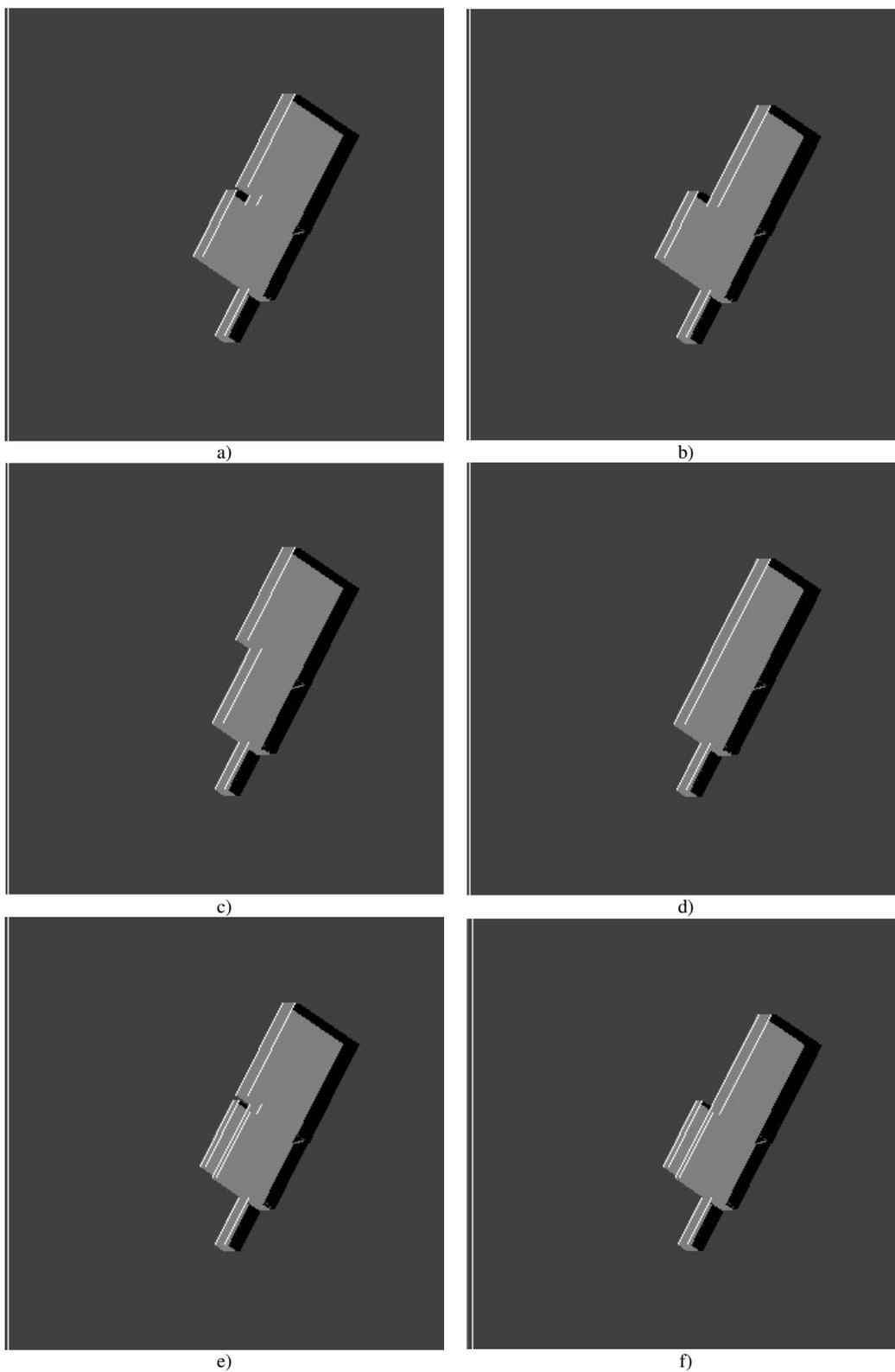


Figure 132 – simulation géométrique des hypothèses générées

Lors du classement des hypothèses, l'hypothèse d est considérée comme hypothèse de référence. Les sections délimitées par des pointillés sont détectées comme des sections surélevées. Les hypothèses contenant ces zones sont donc prioritaires. Les hypothèses a et e contiennent les deux zones donc la sélection est effectuée parmi ces deux hypothèses. L'hypothèse avec un MDL minimum est l'hypothèse a : il s'agit de l'hypothèse la plus proche topologiquement de la vérité terrain.

zone 4 :

La dernière zone testée représente un bâtiment rectangulaire avec une cour intérieure. Le toit de ce bâtiment est plat à une hauteur de 5m (BDTopo^{©IGN}).

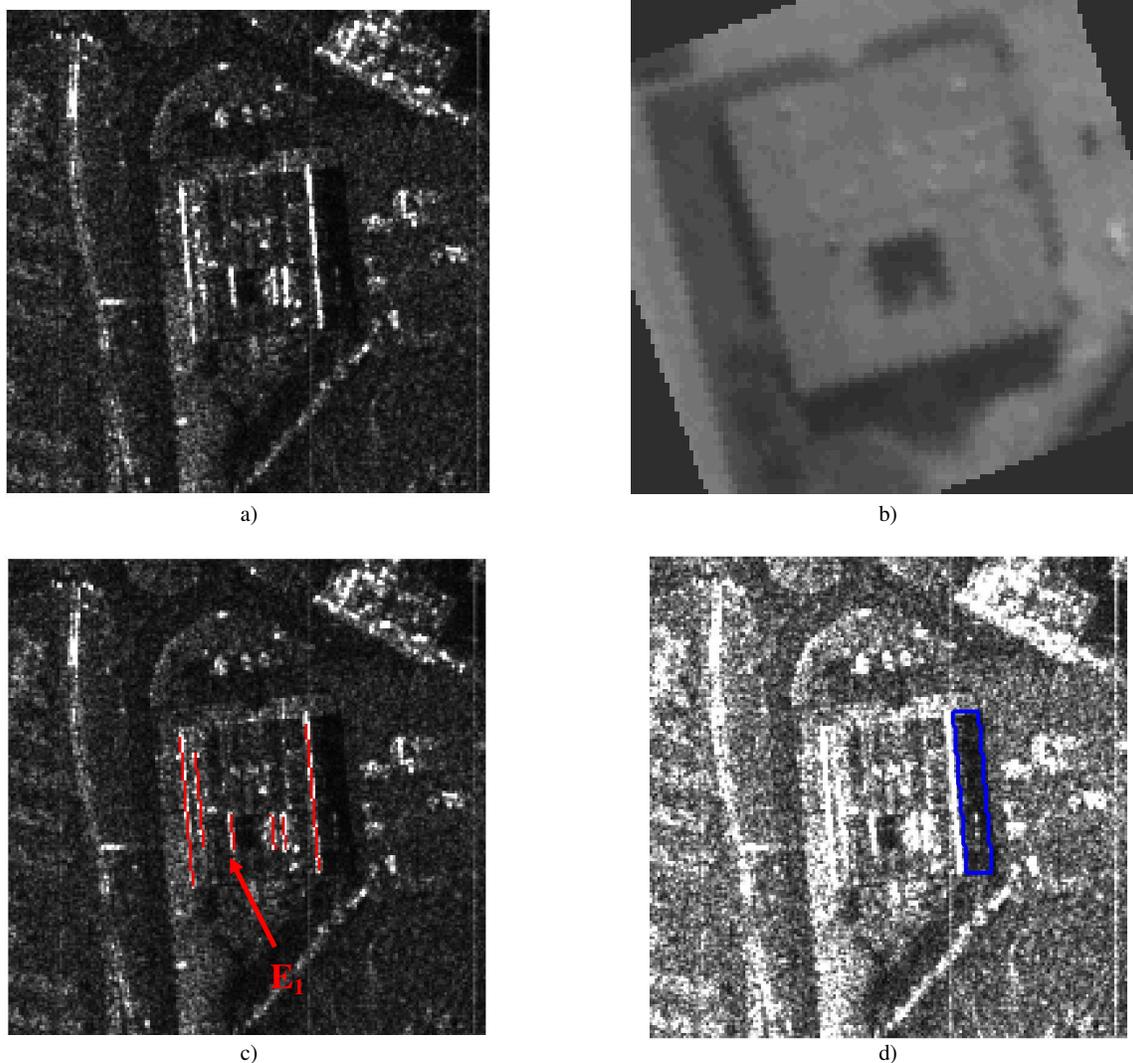
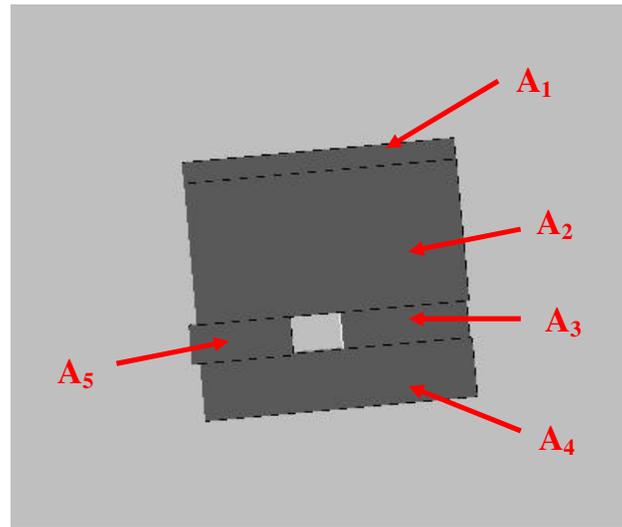


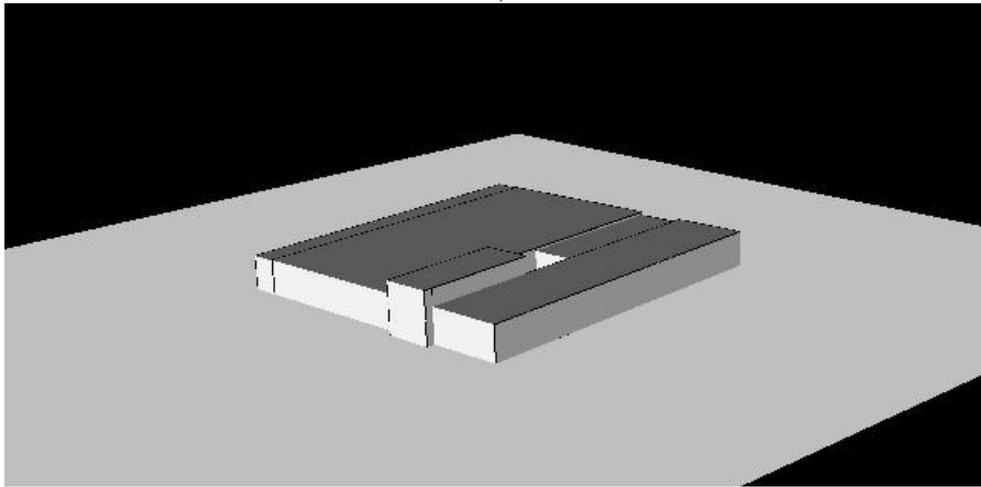
Figure 133 – a) signature RADAR d'un bâtiment, b) image optique correspondante, c) échos et d) ombres extraites

L'algorithme de gestion d'hypothèses fournit une seule hypothèse représentée sous deux angles sur la figure 134. La cour intérieure est reconstruite. Sur les parties A₁ à A₄ la hauteur du toit vaut environ 4.5m. La partie A₅ est à 6.5m. Cette sur estimation de la hauteur provient de la hauteur de l'écho E₁ (cf. figure 133c) qui est, pour une raison encore non expliquée, estimée vers 7m. Le RSB de cet écho étant très important, ce terme est prépondérant lors de la fusion des hauteurs pour la détermination de la hauteur du toit. Néanmoins, une grande partie

du bâtiment est reconstruite avec 50cm d'erreur. De plus, géométriquement le bâtiment reconstruit est cohérent avec la vérité terrain.



a)



b)

Figure 134 – deux vues de l'hypothèse générée

5.8. Conclusion

Dans ce chapitre nous avons proposé une méthode de reconstruction de bâtiment à partir d'un ensemble de primitive incomplet par gestion d'hypothèses. La gestion d'hypothèses consiste à relier les primitives extraites pour satisfaire les hypothèses d'empreintes à angles droits et de toits plats ou plats par morceaux. La validation des hypothèses est effectuée à partir de règles de logique (géométrie, surfaciques). Un algorithme de classement d'hypothèses a également été proposé. Il est non optimal car il utilise un simulateur géométrique ne fournissant pas de données interférométriques. Les signatures obtenues ne peuvent donc pas être comparées efficacement à l'interférogramme. Néanmoins, il permet de fournir à l'utilisateur des indices pour la décision finale. La méthode de reconstruction a ensuite été validée sur images RAMSES haute résolution et a permis de reconstruire des bâtiments de formes diverses avec une précision altimétrique pouvant atteindre 50cm.

Un point fort de la méthode proposée est que la méthode fonctionne même si certains types de primitives ne sont pas extraits : sur chaque bâtiment, seulement deux des trois types de primitives (ombres, échos, textures) sont utilisés. Elle permet d'obtenir une reconstruction géométriquement correcte lorsque le début et la fin du bâtiment sont bien délimités. Dans le cas contraire, la détection des zones élevées sur l'interférogramme semblerait plus adaptée.

La précision altimétrique sur la reconstruction dépend du type de primitives disponibles et de la détection des façades. Lorsque les façades sont détectées, la précision altimétrique obtenue sur la reconstruction est très bonne (50cm). Dans ce cas, la méthode proposée présente une meilleure précision altimétrique que celles obtenues dans [Bolter 2001] à partir de quatre acquisitions orthogonales (1.25m) et dans [Soergel 2003] à partir d'une, voir deux acquisitions orthogonales (1.5m). En revanche, les résultats de la méthode proposée sont moins bons lorsque l'information surfacique (interférogramme) est la seule information de hauteur disponible.

6. Extensions

6.1. Introduction

Dans ce chapitre, les contributions apportées dans cette thèse sont utilisées sur d'autres jeux de données que les images de Dunkerque. Tout d'abord le modèle de mélange interférométrique est étendu puis validé sur données polarimétriques interférométriques d'Avignon. Il est également comparé au modèle de mélange polarimétrique interférométrique proposé dans [Colin 2005]. Ensuite la méthode de reconstruction est testée sur un autre jeu de données interférométriques monopasses en bande X avec une meilleure résolution. Le comportement de la méthode de reconstruction est étudié sur ces images. Enfin, nous proposons des extensions de la méthode de reconstruction, notamment pour affiner les modèles reconstruits et étendre le type de toits des modèles de bâtiments.

6.2. Vers l'interférométrie polarimétrique

La polarimétrie consiste à utiliser la polarisation de l'onde électromagnétique, c'est-à-dire, la courbe décrite par l'extrémité du champ électrique en fonction du temps dans un plan d'onde fixé. Le plan d'onde est un plan perpendiculaire au vecteur d'onde \vec{k} , support de la direction de propagation (cf. figure 135).

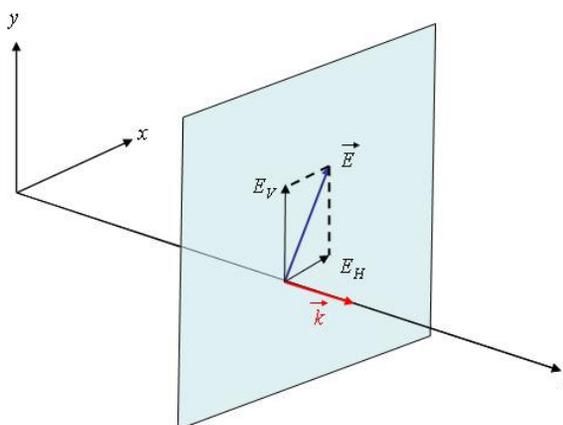


Figure 135 – champ électrique et plan d'onde

L'interférométrie polarimétrique reprend le même principe que l'interférométrie mais utilise la polarisation de l'onde RADAR sur les deux antennes du capteur. Il est alors possible d'étudier plus précisément les mécanismes (le comportement), de l'onde au sein d'une cellule de résolution.

Ce paragraphe présente tout d'abord succinctement l'imagerie RSO polarimétrique et interférométrique polarimétrique. Les données interférométriques polarimétriques peuvent être considérées comme plusieurs interférogrammes obtenus dans des polarisations différentes en émission et réception. Le modèle de mélange interférométrique proposé au §2.5 peut donc être adapté à ces données.

La séparation de mélange est également un thème de recherche en interférométrie polarimétrique. Un modèle de mélange d'au maximum trois contributions ainsi qu'une technique d'optimisation de cohérence, proposés dans [Colin 2005], sont présentés. L'optimisation de cohérence est une technique permettant d'obtenir un interférogramme débruité à partir des données interférométriques polarimétriques.

Les deux modèles de mélanges (interférométrique et interférométrique polarimétrique) sont alors comparés sur l'estimation de hauteur interférométrique de bâtiments. Le modèle de mélange interférométrique est également testé sur l'interférogramme débruité obtenu à l'issue de l'optimisation de cohérence.

6.2.1. Notions de polarimétrie

Le champ électromagnétique dans le plan d'onde peut être représenté dans la base horizontale / verticale (notée h/v). Il s'écrit :

$$\vec{E} = \begin{bmatrix} E_h \\ E_v \end{bmatrix} = \begin{pmatrix} a_h \cos(\omega t) \\ a_v \cos(\omega t + \delta) \end{pmatrix} \quad (92)$$

où a_h et a_v représentent les composantes horizontales et verticales du champ électrique
 ω est la pulsation du champ électrique
 δ le déphasage entre les composantes horizontales et verticales du champ électrique

L'équation de l'extrémité du champ électrique peut alors être transformée sous la forme d'une ellipse :

$$\left(\frac{E_h}{a_h}\right)^2 + \left(\frac{E_v}{a_v}\right)^2 - \frac{2E_h E_v \cos(\delta)}{a_h a_v} = \sin^2(\delta) \quad (93)$$

La polarisation est dite linéaire lorsque le déphasage entre les composantes horizontales et verticales de l'ellipse vaut 0 [π]. L'ellipse se réduit à un segment. Une polarisation circulaire est obtenue lorsque $\delta = 45^\circ[\pi]$.

L'état de polarisation, ou vecteur de Jones, est défini par :

$$\vec{P} = \begin{pmatrix} A_h \\ A_v \end{pmatrix} = \begin{pmatrix} a_h \\ a_v \exp(j\delta) \end{pmatrix} = \begin{pmatrix} a \cos(\alpha) \\ a \sin(\alpha) \exp(j\delta) \end{pmatrix} \quad (94)$$

où α représente l'inclinaison de l'ellipse représentée sur la figure 136.

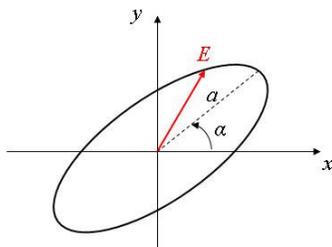


Figure 136 – champ électrique en polarisation elliptique

La matrice de Sinclair, encore appelée matrice de diffusion, est notée S et permet de relier les vecteurs de Jones des ondes émises et reçues.

$$S = \begin{pmatrix} S_{hH} & S_{hV} \\ S_{vH} & S_{vV} \end{pmatrix} \quad (95)$$

Les notations x et X représentent respectivement les polarisations des ondes émises et reçues. Dans les conditions d'acquisition des images utilisées (monopasse, bande X), le théorème de réciprocité de Lorentz est valide et permet d'écrire : $S_{hV} = S_{vH}$.

La matrice de cohérence T de l'onde s'exprime par $\langle \vec{P}\vec{P}^\dagger \rangle$ où \vec{P} est le vecteur de Jones, † représente le conjugué transposé et $\langle \rangle$ le produit scalaire. Elle permet d'obtenir une description complète des propriétés diffusantes de la cible et s'écrit :

$$T = \begin{pmatrix} E(A_h A_h^*) & E(A_h A_v^*) \\ E(A_v A_h^*) & E(A_v A_v^*) \end{pmatrix} \quad (96)$$

où $E(\)$ désigne l'espérance mathématique.

Le degré de cohérence polarimétrique est défini par :

$$\gamma = \frac{E(A_h A_v^*)}{E(A_h A_h^*) E(A_v A_v^*)} \quad (97)$$

En polarimétrie, la polarisation de l'onde est prise en compte pour obtenir des informations supplémentaires sur la scène telles que les types de mécanismes subis par l'onde au sein d'une cellule de résolution. Pour cela, la matrice de diffusion est projetée dans une base de matrice pour former le vecteur cible. La base choisie dans ce paragraphe est la base de Pauli [Cloude 1996].

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

En effet les deux premières matrices de cette base représentent des matrices de diffusions d'éléments canoniques connus : sphères ou plaques (échos simples), dièdres (échos doubles). Elle est donc bien adaptée à notre modèle de mélange interférométrique où la distinction entre échos simples et doubles est importante.

Le vecteur de cible s'écrit alors :

$$k_{\vec{P}} = \frac{1}{\sqrt{2}} \begin{pmatrix} S_{hH} + S_{vV} \\ S_{hH} - S_{vV} \\ 2S_{hV} \end{pmatrix} \quad (98)$$

où $S_{hH} + S_{vV}$ décrit les phénomènes avec un nombre de rebonds impair (simples et triples...)

$S_{hH} - S_{vV}$ décrit les phénomènes avec un nombre de rebonds pair (doubles)

$2S_{hV}$ décrit les rétrodiffusions volumiques qui apparaissent notamment dans des applications de pénétration de feuillages.

6.2.2. L'interférométrie polarimétrique

En interférométrie, la polarisation de l'onde RADAR peut également être utilisée pour obtenir des données interférométriques polarimétriques. Pour chaque canal d'émission-réception, un interférogramme est formé. Trois interférogrammes sont donc à disposition (les interférogrammes provenant des canaux hV et vH sont considérés comme identiques).

La matrice de cohérence généralisée est introduite par [Papathanassiou 1997] et [Cloude 1998] comme la matrice bloc :

$$T = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \quad (99)$$

Avec $T_{ij} = \langle \vec{k}_{p,i} \vec{k}_{p,j}^\dagger \rangle$ et $\vec{k}_{p,i}$ le vecteur cible pour l'antenne i .

Soient $\vec{\omega}_1$ et $\vec{\omega}_2$ les combinaisons linéaires de polarisations en émission / réception que l'on choisit pour les deux antennes de l'acquisition interférométrique. On appelle $\vec{\omega}_1$ et $\vec{\omega}_2$ des « mécanismes », ils se représentent sous forme d'une matrice normalisée à 3 lignes et 1 colonne. En effectuant les produits matriciels $\vec{\omega}_i \vec{k}_p^i$, $i = \{1, 2\}$, cela revient à étudier uniquement le mécanisme sur l'onde reçue par l'antenne i .

La cohérence généralisée pour ces deux mécanismes est définie par :

$$\gamma(\omega_1, \omega_2) = \frac{\vec{\omega}_1^\dagger T_{12} \vec{\omega}_2}{\sqrt{\vec{\omega}_1^\dagger T_{11} \vec{\omega}_1} \sqrt{\vec{\omega}_2^\dagger T_{22} \vec{\omega}_2}} \quad (100)$$

Le module et l'argument de (100) représentent respectivement le niveau de cohérence et la phase interférométrique de la zone sur laquelle elle est calculée. Selon les mécanismes choisis, la cohérence et la phase interférométrique sont donc différentes. L'ensemble de ces cohérences complexes obtenues lorsque les mécanismes varient est appelé champ de valeur de la cohérence et peut être représenté dans le disque du plan complexe, de rayon 1 et de centre (0,0).

Dans [Colin 2005], l'auteur remarque que l'optimisation de cohérence effectuée par Cloude et Papathanassiou à partir de (100) aboutit à obtenir des mécanismes qui ont sensiblement des composantes de modules égaux. Etant donné qu'en interférométrie, la même cellule de résolution est considérée par les deux antennes à des angles d'incidences très proches, l'auteur propose de choisir $\vec{\omega}_1 = \vec{\omega}_2 = \vec{\omega}$. De cette façon, les estimations des mécanismes par le processus d'optimisation sont numériquement plus stables. La cohérence généralisée s'exprime alors par :

$$\gamma(\omega) = \frac{\vec{\omega}^\dagger T_{12} \vec{\omega}}{\sqrt{\vec{\omega}^\dagger T_{11} \vec{\omega}} \sqrt{\vec{\omega}^\dagger T_{22} \vec{\omega}}} \quad (101)$$

Supposons qu'il n'y ait qu'un rétrodiffuseur dans la cellule de résolution. Dans le cas idéal non bruité, le champ de cohérence se réduirait à un point. Pratiquement, à cause du bruit sur les mesures, le champ de cohérence est plus étalé et est représenté en rouge sur la figure 137. L'étalement du champ de cohérence décroît avec le rapport signal à bruit.

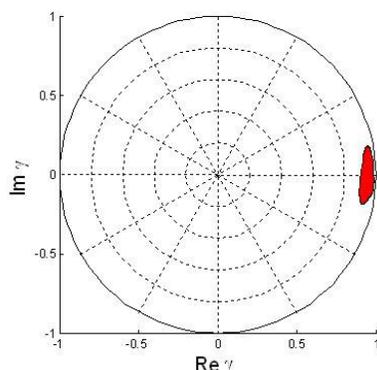


Figure 137 – champ de valeur de la cohérence généralisée pour un unique mécanisme peu bruité

6.2.3. Modèle de mélange de deux contributions ayant des mécanismes différents

Dans une même cellule de résolution, il y a généralement plusieurs diffuseurs et par conséquent, plusieurs mécanismes. Le modèle proposé par [Colin 2005] peut séparer jusqu'à trois contributions suivant des mécanismes différents. Nous nous restreignons ici à des mélanges de deux contributions. Il est montré dans [Chien 1998] que, dans ce cas, la forme du champ de vecteur de la cohérence généralisée est une ellipse (cf. figure 138). Cette forme provient notamment des corrélations entre les différents mécanismes.

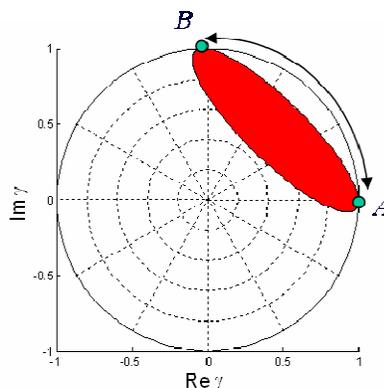


Figure 138 – champ de valeur de la cohérence généralisée pour deux mécanismes différents

En supposant que les mécanismes, considérés séparément, ont une cohérence égale à 1, deux points de l'ellipse, notés A et B , sont obtenus. Les phases de A et B sont respectivement notées φ_A et φ_B . L'angle interférométrique entre A et B permet d'estimer la hauteur du bâtiment notée h_{pol} , par :

$$h_{pol} = \frac{h_{amb} (\varphi_B - \varphi_A)}{2\pi} \quad (102)$$

où h_{amb} est la hauteur d'ambiguïté.

En pratique, l'hypothèse de cohérence égale à 1 n'est pas parfaitement vérifiée mais la cohérence moyenne des images RAMSES est très élevée car l'acquisition est monopasse. L'hypothèse d'absence de décorrélacion temporelle est bien respectée.

6.2.4. Optimisation de cohérence

Le principe de l'optimisation de cohérence est de déterminer la cohérence maximale du champ de valeur (le point M sur la figure 139). Elle est généralement effectuée par la maximisation du numérateur de (101) à dénominateur constant [Cloude 1998]. La hauteur associée à ce maximum de cohérence est alors moins bruitée.

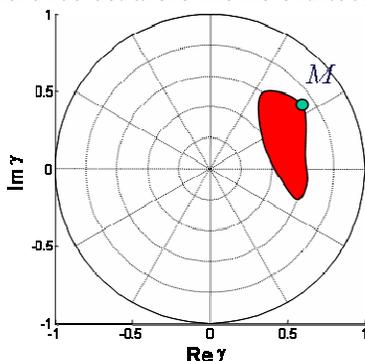


Figure 139 – Point du champ de valeur de cohérence maximale

6.2.5. Estimation de la hauteur de bâtiments sur données interférométriques polarimétriques

6.2.5.1. Présentation des données

La comparaison des modèles de mélanges polarimétriques et interférométriques est effectuée sur des images interférométriques pleinement polarimétriques haute résolution (70cm) issues du capteur RAMSES. L'acquisition a été effectuée sur le site d'Avignon en bande X.

Quelques signatures de bâtiments utilisées sont présentées sur la figure 140. Les images polarimétriques sont représentées dans la base de Pauli.

- rouge pour la composante $hH+vV$
- vert pour la composante $hH-vV$
- bleu pour la composante $2hV$

Sur ces images, les mécanismes surfaciques tels que la rétrodiffusion du sol ou des toits ont visiblement le rouge comme composante dominante. Grâce aux images polarimétriques, l'interprétation de la scène est nettement facilitée.

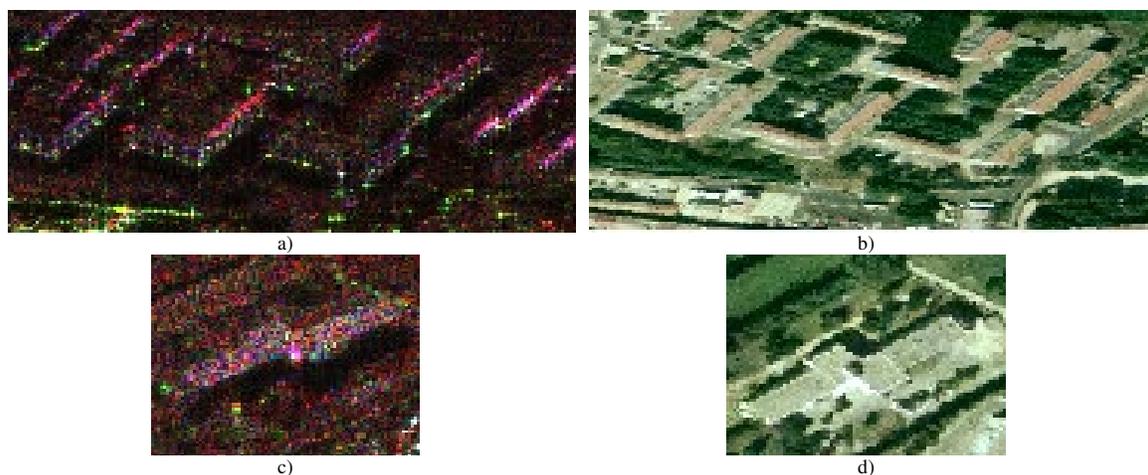


Figure 140 – a) et c) bâtiments sur images polarimétriques, b) et d) images optiques correspondantes

Seule une dizaine de bâtiments, présentant une réponse polarimétrique interprétable, ont été sélectionnés. Des exemples de bâtiment non sélectionnés pour la comparaison des modèles sont présentés sur la figure 141. Les échos doubles, très brillants apparaissent en blanc et non en vert. Cela signifie que les trois mécanismes sont mélangés. L'origine de ce mélange n'est pas encore expliquée. Le modèle interférométrique ne peut pas être utilisé si trois mécanismes sont présents.

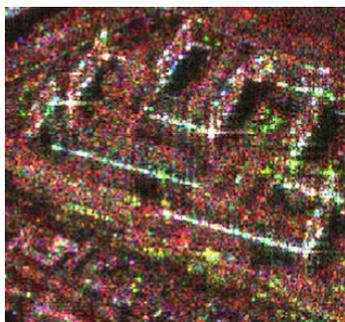


Figure 141 – bâtiments avec des signatures polarimétriques difficiles à interpréter

La vérité terrain était disponible pour quelques bâtiments de la zone. Pour les autres, les mesures monoscopiques ont été utilisées pour estimer de façon précise la hauteur des bâtiments.

6.2.5.2. Comparaison des modèles de mélange interférométriques et polarimétriques

Les modèles sont comparés sur l'estimation de hauteur des bâtiments. Au début de la zone de repliement, seulement deux contributions sont mélangées. Nous nous plaçons dans les conditions d'application des modèles interférométriques et polarimétriques.

Comme les canaux de polarisations croisées fournissent des données considérées similaires, les données interférométriques polarimétriques peuvent être considérées comme trois jeux de données interférométriques de la même scène.

Le modèle de mélange interférométrique peut donc être appliqué séparément sur chaque interférogramme pour obtenir trois estimations de hauteurs de bâtiments. La hauteur finale du bâtiment consiste alors en la moyenne des hauteurs corrigées par le modèle. Les moyennes et écart-types du rapport moyen des hauteurs interférométriques et monoscopiques sont consignés dans le tableau ci-dessous.

	Rapport moyen des hauteurs interférométriques et monoscopiques		Ecart-type sur le rapport des hauteurs interférométriques et monoscopiques	
	sans modèle	avec modèle	sans modèle	avec modèle
Polarisation hH	0.85	0.98	0.12	0.09
Polarisation hV	0.75	0.89	0.16	0.09
Polarisation vV	0.87	0.99	0.15	0.13
Moyenne sur les trois polarisations	0.82	0.95	0.12	0.08
Moyenne sur les polarisations hH et vV	0.86	0.98	0.12	0.09

Les erreurs quadratiques moyennes (EQM) pour ces mêmes polarisations sont présentées dans le tableau suivant.

	EQM (m)	
	sans modèle	avec modèle
Polarisation HH	1.41	0.53
Polarisation HV	1.90	0.92
Polarisation VV	1.29	0.77
Moyenne sur les trois polarisations	1.43	0.57
Moyenne sur les polarisations hH et vV	1.29	0.53

Sans modèle interférométrique, le rapport moyen des hauteurs interférométriques et monoscopiques varie, sur chaque polarisation, entre 0.75 et 0.87. Sur les 3 interférogrammes,

la mesure interférométrique sous-estime la hauteur réelle du bâtiment. L'utilisation du modèle interférométrique permet d'améliorer ce rapport et donc l'estimation de hauteur. Pour les polarisations hH et vV ce rapport atteint respectivement 0.98 et 0.99. Sur ces deux polarisations, grâce au modèle interférométrique, les hauteurs de bâtiments sont bien estimées.

Même si, sur la polarisation croisée, le modèle interférométrique permet d'augmenter le rapport, il reste néanmoins faible. Ceci provient des primitives utilisées pour l'estimation de hauteur. En effet, les hauteurs de bâtiments sont estimées à partir des hauteurs des échos simples et des échos doubles. Ces deux hauteurs peuvent être discriminées polarimétriquement grâce à la décomposition dans la base de Pauli sur les composantes hH + vV et hH - vV. Pour ces primitives, les informations contenues dans la composante vH de la base sont donc très bruitées, d'où une mauvaise estimation de la hauteur des bâtiments avec et sans modèle sur ce canal.

Lorsqu'on considère la hauteur moyenne sur les deux polarisations adaptées à l'estimation de hauteur des bâtiments, le rapport de hauteur atteint 0.98. De plus l'écart type sur le rapport des hauteurs diminue.

Lorsque les erreurs quadratiques moyennes avec et sans modèles sont comparées, il est visible que le modèle interférométrique permet de diminuer de façon importante l'erreur quadratique moyenne commise sur l'estimation de hauteur interférométrique. Elle passe de 1.3m à 0.5m lorsque les deux polarisations adaptées sont utilisées.

Sur le tableau ci-dessous sont présentés les résultats d'estimation de hauteurs grâce au modèle polarimétrique. Les résultats trouvés sont similaires à ceux proposés par le modèle interférométrique moyenné sur les polarisations hH et vV.

Rapport moyen des hauteurs interférométriques et monoscopiques		Ecart-type sur le rapport des hauteurs interférométriques et monoscopiques		EQM (m)	
0.99		0.09		0.56	

Les deux algorithmes de séparation de mélanges de deux contributions sont donc validés sur ces données.

6.2.5.3. Utilisation du modèle de mélange interférométrique sur l'interférogramme optimisé

L'optimisation de cohérence de [Colin 2005] est effectuée sur les mêmes images. Les résultats sur l'estimation de hauteurs sur l'interférogramme optimisé sont présentés dans le tableau ci-dessous avec et sans le modèle interférométrique.

Rapport moyen des hauteurs interférométriques et monoscopiques		Ecart-type sur le rapport des hauteurs interférométriques et monoscopiques		EQM (m)	
sans modèle	avec modèle	sans modèle	avec modèle	sans modèle	avec modèle
0.93	1.09	0.14	0.14	1.04	1.09

Il est visible que l'optimisation de cohérence permet d'améliorer l'estimation de hauteur par rapport aux données non optimisées. Le rapport moyen des hauteurs interférométriques et monoscopiques passe de 0.82 (en moyenne sur les trois polarisations) à 0.93. En revanche l'écart-type augmente légèrement de 0.12 à 0.14. Lorsque le modèle interférométrique est utilisé, le rapport passe à 1.09 : la hauteur des façades est surestimée. En effet, lors de la maximisation de cohérence, le mélange altimétrique entre le sol et le bord du toit est séparé lorsque le maximum de cohérence correspond à la contribution de toit. Dans les autres cas, l'optimisation de cohérence donne une hauteur qui ne correspond pas à la hauteur du

bâtiment. Que le mélange soit séparé ou non, l'utilisation du modèle interférométrique n'est plus valide car l'information de mélange est perdue. C'est pourquoi l'optimisation de cohérence ne devrait être utilisée que sur les zones sans repliement.

6.2.6. Conclusion

Les modèles de mélange interférométrique appliqué aux données interférométriques polarimétriques et mélanges polarimétriques ont été validés sur l'estimation de hauteur de bâtiments sur images RAMSES haute résolution. Ils améliorent nettement l'estimation de hauteur des bâtiments pour atteindre une erreur quadratique moyenne de l'ordre de 50cm pour une hauteur d'ambiguïté moyenne de 170m. L'optimisation de cohérence permet de débruiter l'interférogramme mais, dans les zones de repliement, perd l'information de mélange. Il doit être utilisé seulement sur les zones sans repliement.

6.3. Vers la très haute résolution

Dans ce paragraphe, nous voulons vérifier que l'algorithme de reconstruction proposé est toujours valide sur un autre jeu de données lorsque la résolution augmente. Les conséquences principales de l'affinement de la résolution sur la méthode de reconstruction sont tout d'abord détaillées. Ensuite le nouveau jeu de données, utilisé dans ce paragraphe, est présenté. Enfin, les résultats de la méthode de reconstruction sont présentés et analysés.

6.3.1. Influence de l'amélioration de la résolution

L'augmentation de la résolution a des conséquences géométriques (cf. §2.2.2, p31), radiométriques et interférométriques.

Lorsque la résolution s'améliore, la validité de l'hypothèse d'onde plane et donc la mesure monoscopique telle qu'elle est présentée au §2.3.1.2 p38 n'est plus forcément valide. La hauteur monoscopique doit alors être calculée avec les angles d'incidences locaux.

Affiner la résolution revient à voir plus de détails sur la zone d'étude. Par conséquent, le nombre de primitives (échos, ombres) à extraire sur l'image d'amplitude devrait être plus important. De plus, l'augmentation de la résolution selon l'axe des distances implique d'utiliser une bande utile plus importante. Un plus large spectre de fréquences est donc utilisé et les textures devraient également être plus fréquentes.

La dynamique entre les réflexions ponctuelles et les rétrodiffusions surfaciques augmente également lorsque la résolution augmente. Ce phénomène vient du traitement RSO. En effet, considérons un unique réflecteur dans la cellule de résolution et supposons que le signal soit utilisé N fois dans la bande utile et la bande doppler. Cela revient à considérer le même réflecteur N^2 fois dans la sommation cohérente. Lorsque la compression et la synthèse d'ouverture sont effectuées parfaitement, toutes ces contributions ont la même phase propre donc l'amplitude résultant est N^2 fois l'amplitude de la même contribution. Maintenant considérons une rétrodiffusion surfacique ayant lieu dans une cellule de résolution similaire. Chaque réflecteur de la cellule de résolution possède une phase propre différente. Par conséquent l'amplitude résultante est obtenue, par un moyennage quadratique. L'amplitude résultante correspond à seulement N fois la même contribution. L'amplitude des réflexions spéculaires augmente donc plus vite que celle des rétrodiffusions surfaciques.

Enfin, pour améliorer la résolution en azimuth avec une même antenne, il faut augmenter le temps d'intégration lors de la synthèse d'ouverture. La partie de la trajectoire du porteur sur laquelle la synthèse est effectuée devient alors plus importante. La compensation de

mouvement se révèle alors cruciale pour obtenir un signal correctement focalisé. Elle doit être effectuée précisément pour les deux antennes d'acquisition du signal interférométrique.

6.3.2. Présentation des données

Les données utilisées dans ce paragraphe sont des images interférométriques de meilleures résolutions que celles de Dunkerque. Pour des raisons de confidentialité, nous ne pouvons préciser ces résolutions. La chaîne de reconstruction a été testée sur deux bâtiments. La vérité terrain sur la hauteur de ces bâtiments n'était pas disponible. Seule une évaluation topologique sera possible.

Le premier bâtiment, dont la signature est présentée sur la figure 142a, est un bâtiment rectangulaire. La signature du toit est constituée d'un ensemble d'écho dans l'orientation secondaire. Ces échos proviennent de la structure du toit en dents de scie (cf. figure 142b).

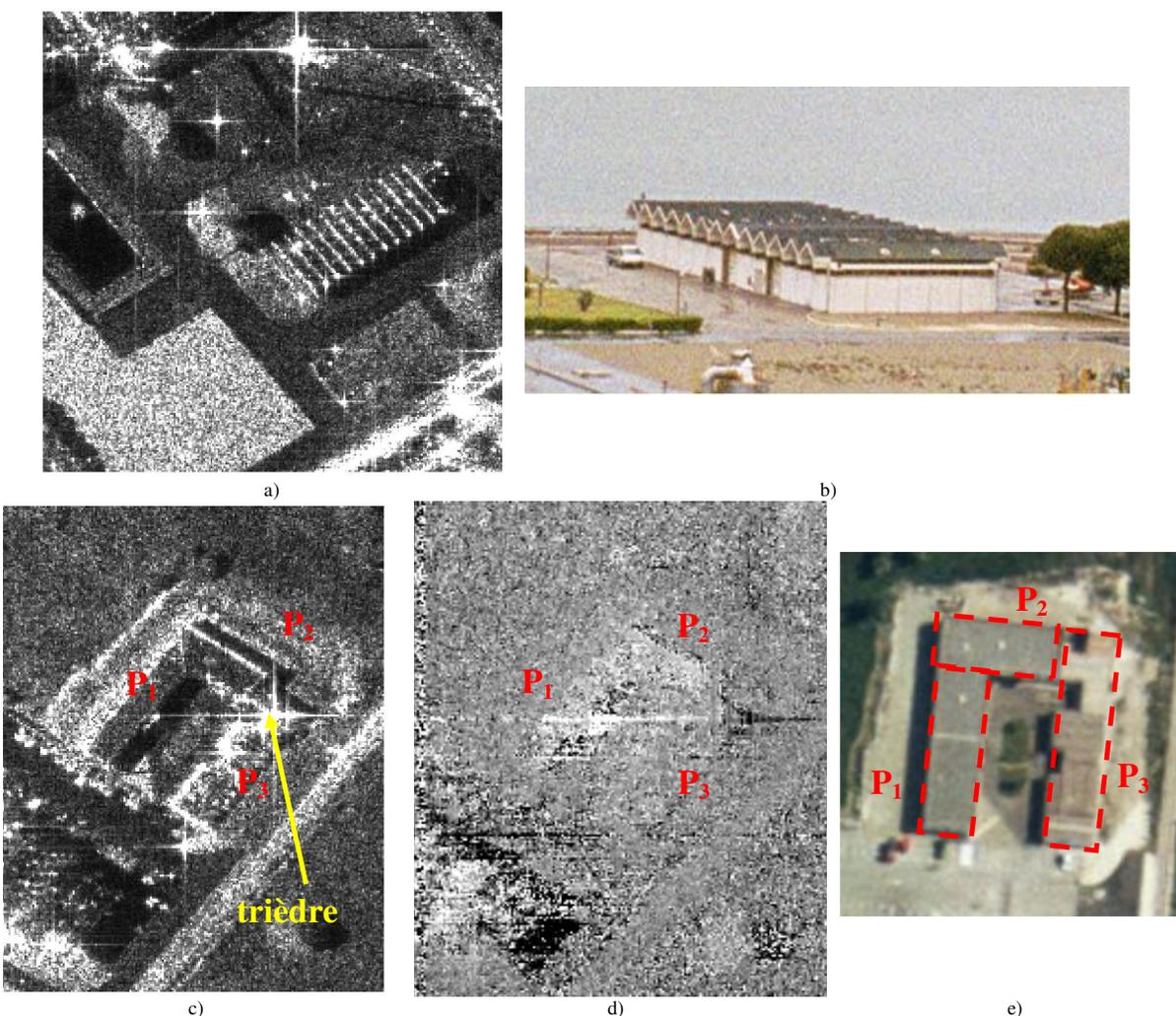


Figure 142 – a), c) signatures de bâtiments sur images RSO,
d) interférogramme correspondant à c : les zones élevées apparaissent en gris clair par rapport au sol (en gris foncé)
b) et e) images optiques de a et c

Le deuxième bâtiment est un bâtiment en *U* que l'on peut approcher par trois parallélépipèdes rectangles, notés P_1 , P_2 , P_3 , et représentés en traits discontinus rouges sur l'image optique correspondante (figure 142e). La signature de ce bâtiment est présentée sur la

figure 142c. Les toits de P_1 et P_2 sont texturés. Sur P_3 , des échos doubles sont visibles. Le trièdre, formé par le sol et les façades de P_2 et P_3 est également visible. Au regard de la taille de l'ombre, P_3 semble peu élevé. Ceci est confirmé sur l'interférogramme (cf. figure 142d) où P_3 est moins haut que P_1 et P_2 et également sur l'image optique où l'ombre projetée par P_3 est plus petite que celles projetées par P_1 et P_2 .

6.3.3. Résultats de la méthode de reconstruction

zone a :

Sur la figure 143a, l'ombre détectée par croissance de région après fermeture est représentée en bleu. Les contours entre l'ombre et le sol sont bien délimités. Ceux entre le bâtiment et l'ombre le sont moins car, sur les parties où il n'y a pas d'échos forts, le toit est peu rétrodiffusant. Grâce à la régularisation (cf. figure 143b), le bruit sur la détection est en partie éliminé.

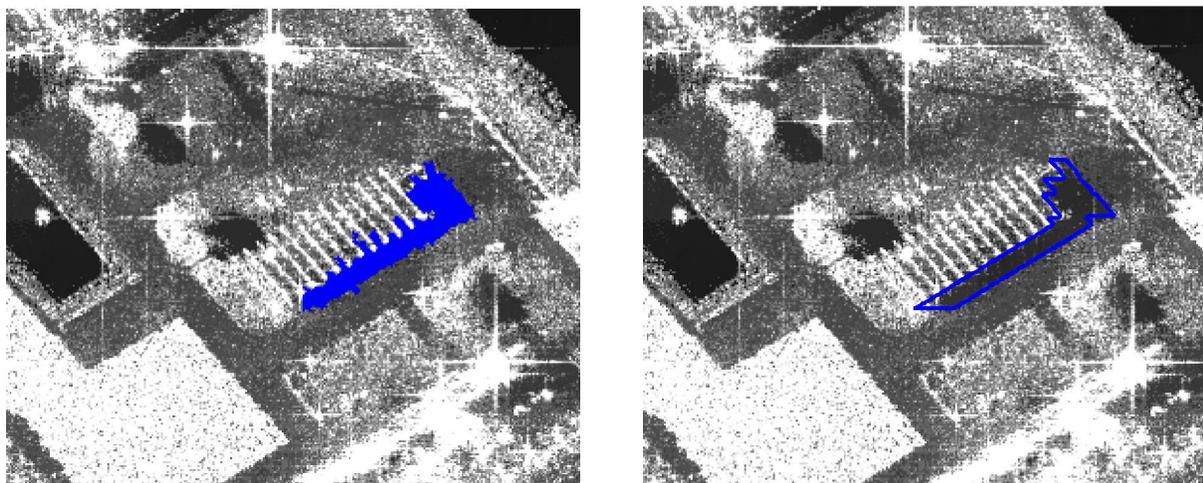


Figure 143 – a) ombre détectée par croissance de région et fermeture, b) extraction après régularisation avec $\Lambda_0=5$ et $\Lambda_1=5$

Les échos forts du toit correspondant aux creux des dents de scie sont correctement extraits et sont représentés sur la figure 144

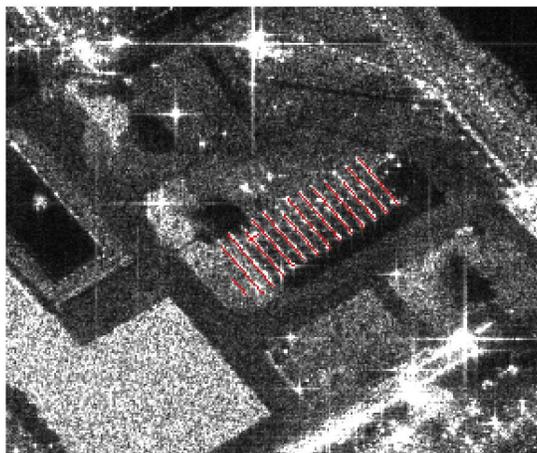


Figure 144 – résultats de la croissance de rectangle pour des paramètres $\beta=1^\circ$, $\varepsilon=1$ pixel, variation de i_0 et j_0 de ± 1 pixels taille mini = 2m

A l'issue de l'étape de reconstruction, une seule hypothèse, présentée sur la figure 145, est obtenue. L'empreinte au sol du bâtiment est globalement rectangulaire. La structure du toit ne pouvait pas être retrouvée correctement car les toits en pente ne sont pas modélisés. La hauteur du bâtiment est estimée entre 5.5m et 6m. Ce qui semble cohérent avec l'image optique.

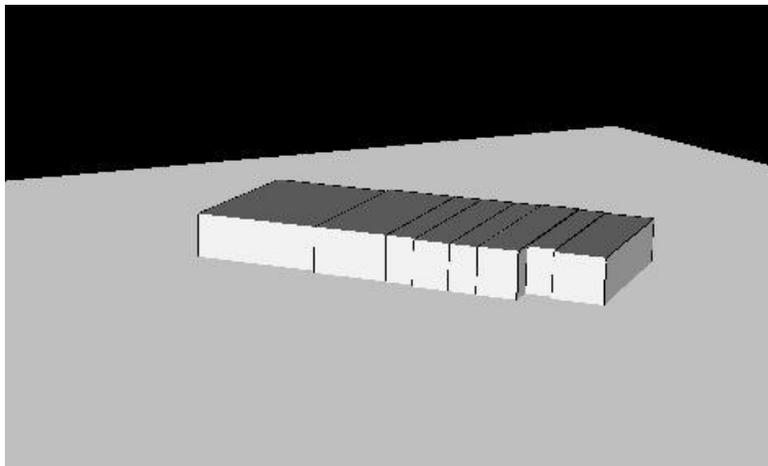


Figure 145 – hypothèse générée

zone b :

Les ombres détectées et extraites sont représentées sur la figure 146. L'ombre projetée par le bâtiment 1 est correctement extraite. L'ombre O_2 est projetée par P_2 et P_3 . L'ombre projetée par P_2 est bien extraite. Celle correspondant à P_3 est légèrement trop large. P_3 projette également une ombre selon l'orientation principale (représentée par la flèche rouge) et n'a pas été détectée car elle est très petite.

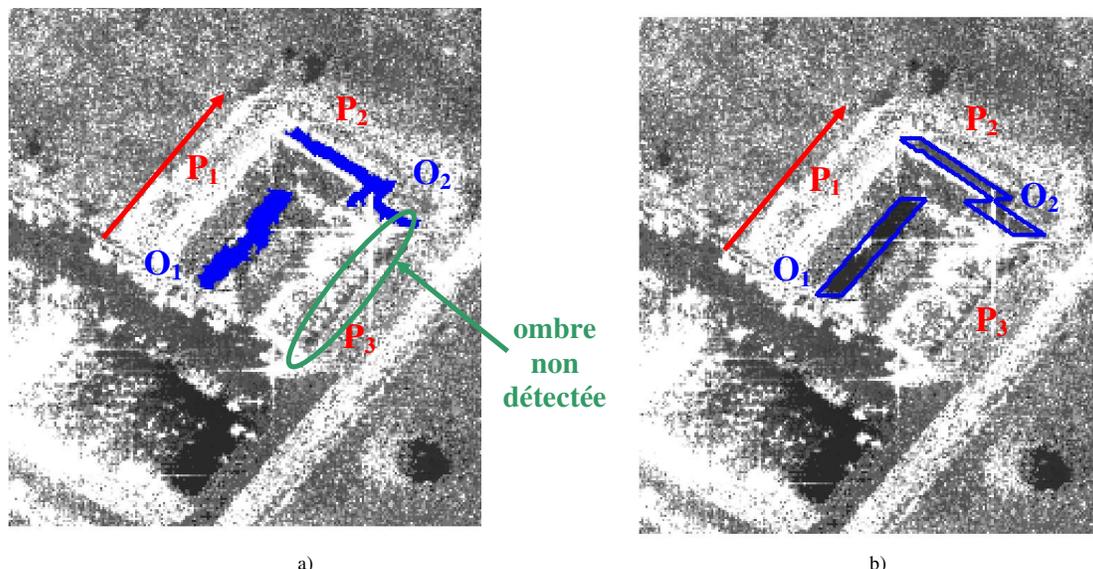


Figure 146 – a) ombres détectées par croissance de région et fermeture, b) extraction après régularisation avec $\Lambda_0=3$ et $\Lambda_1=10$

Les textures de P_1 et P_2 sont bien détectées (cf. figure 147a). L'algorithme d'extraction fournit un bon résultat, correspondant aux toits de P_1 et P_2 , présenté sur la figure 147b.

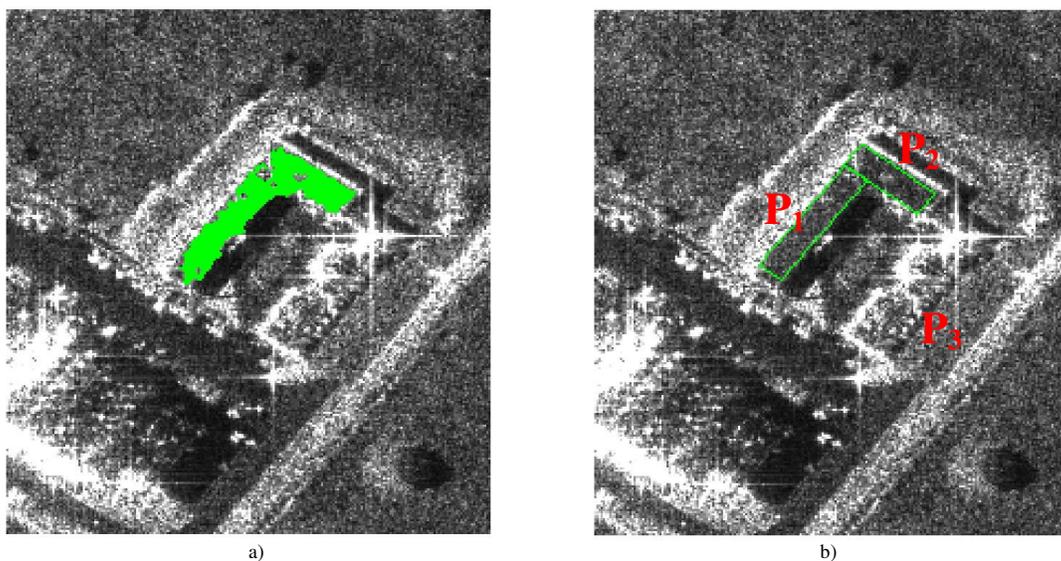


Figure 147 – a) texture détectée par croissance de région et fermeture, b) extraction après régularisation avec $\Lambda_r=3$ et $\Lambda_t=10$

Les échos extraits correspondant principalement à la partie P_3 du bâtiment sont présentés sur la figure 148.

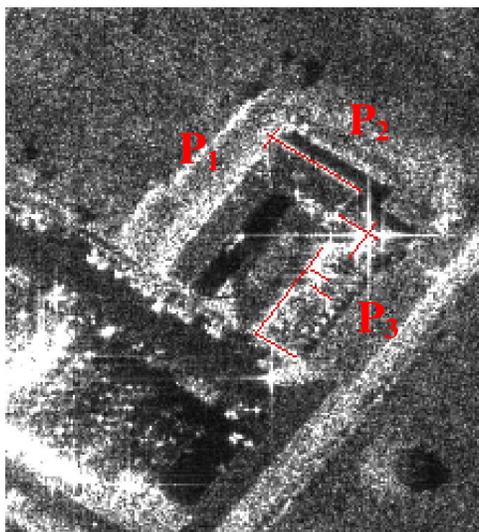


Figure 148 – résultats de la croissance de rectangle pour des paramètres $\beta=1^\circ$, $\varepsilon=1$ pixel, variation de i_0 et j_0 de ± 1 pixels taille mini = 2m

Lorsque seules les ombres et les textures sont utilisées, le bâtiment est reconstruit en L (cf. figure 149). En effet, sans les échos, trop peu d'informations sont disponibles sur la partie P_3 du bâtiment et elle n'est pas reconstruite.

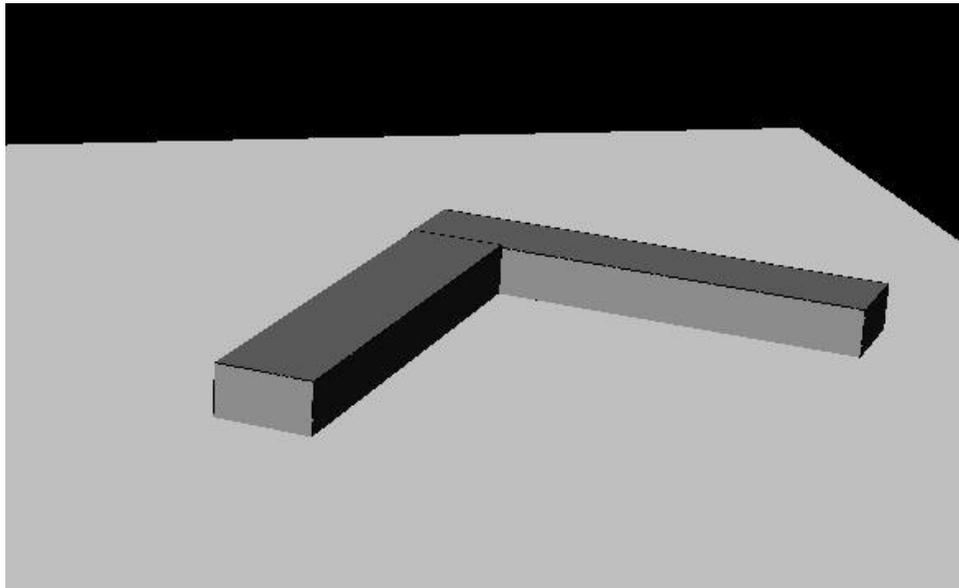


Figure 149 – hypothèse générée à partir des ombres et des textures

Lorsque les échos forts sont également utilisés, une vue du dessus des deux hypothèses générées avec échos forts est présentée sur la figure 150. La seule différence entre ces deux hypothèses est entourée en rouge sur la figure 150a. L'hypothèse minimisant la complexité stochastique est l'hypothèse de la figure 150b. La zone entourée en rouge présente une hauteur interférométrique de 2.5m. La méthode de classement proposée au chapitre 5 privilégie donc l'hypothèse de la figure 150a.

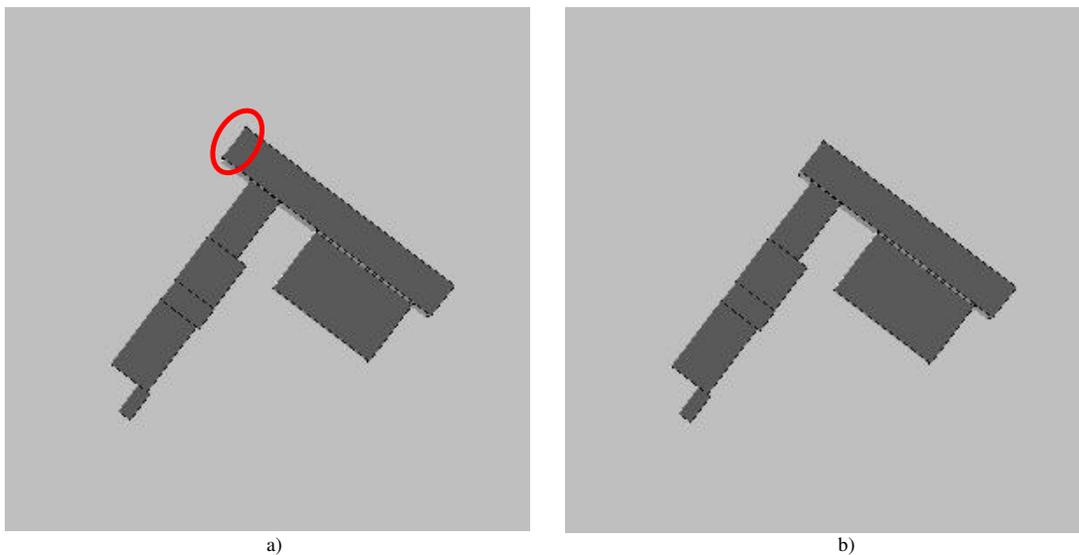


Figure 150 – hypothèses générées à partir des ombres, textures et des échos

Une autre vue de l'hypothèse de la figure 150a est présentée ci-dessous :

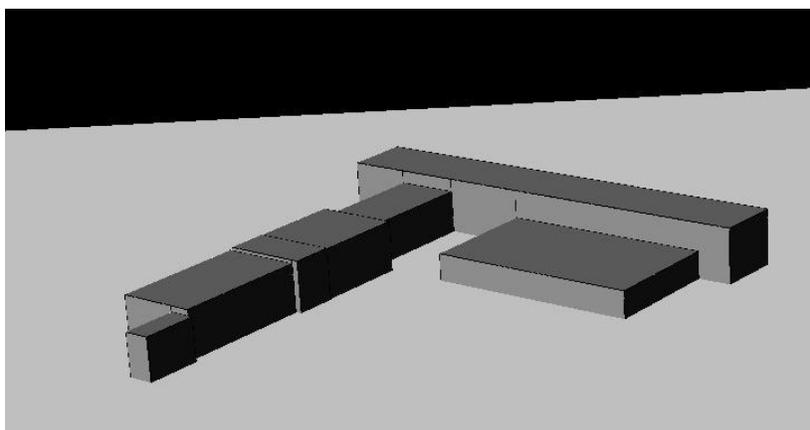


Figure 151 – hypothèse classée numéro 1

Cette hypothèse est maintenant comparée à celle obtenue lorsque les échos forts ne sont pas utilisés. Ces deux hypothèses sont présentées sur la figure 152. Dans les deux cas, P_1 et P_2 sont reconstruits. Une partie de P_3 , délimitée en traits discontinus bleus, est également obtenue.

Lorsque les échos forts sont utilisés (figure 152b), P_1 est découpé en plusieurs zones. Sur chaque zone l'estimation surfacique de hauteur, effectuée par l'algorithme, n'aboutit pas exactement à la même hauteur de toit. C'est pourquoi la reconstruction de P_1 est plus bruitée. Néanmoins, les échos forts permettent de reconstruire une autre partie de P_3 .

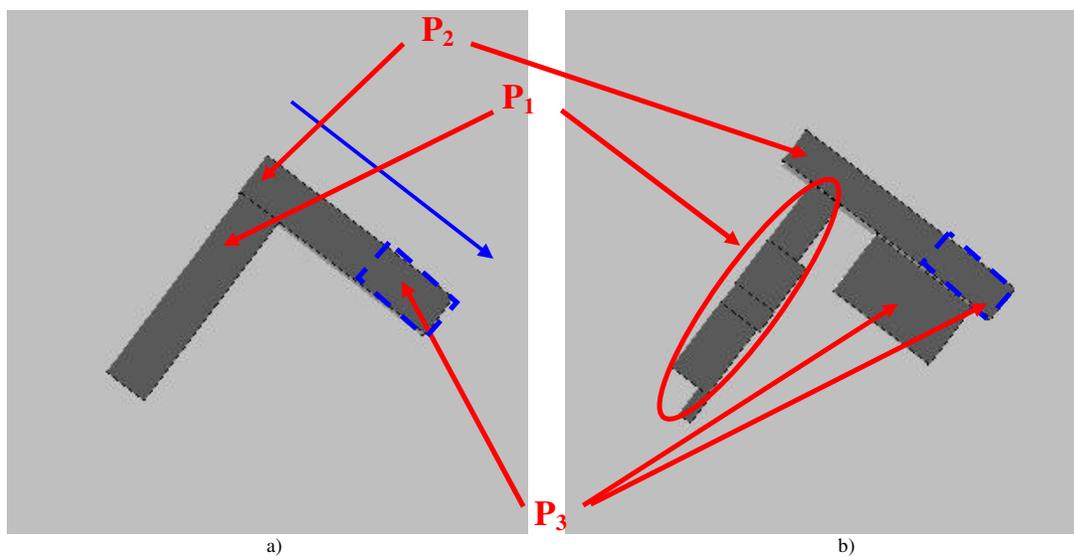


Figure 152 – hypothèse générée a) sans les échos forts et b) avec les échos

Lorsque les échos ne sont pas utilisés, les hauteurs de P_1 et P_2/P_3 valent respectivement 4.69m et 3.78m. Avec les échos forts, la hauteur de P_1 varie entre 3.5m et 4,8m. P_2/P_3 est à 5.2m et la deuxième partie de P_3 est à 2.56m. La hauteur de P_2 est plus élevée car un écho fort apparaît le long de P_2 et fournit une information altimétrique supplémentaire.

Au niveau de la planimétrie, P_1 et P_2 semblent bien reconstruit mais plus fins que sur l'image optique. La partie de P_3 reconstruite lorsque les échos forts sont utilisés semble trop longue selon l'orientation secondaire représentée par la flèche bleue sur la figure 152a. Le trop peu d'informations ne permet pas de délimiter correctement cette partie de bâtiment.

6.3.4. Conclusion

La méthode de reconstruction a été validée sur des images de meilleures résolutions sur deux bâtiments. Dans le premier cas, l’empreinte du bâtiment est bien reconstruite mais pas le toit car il ne correspond pas aux modèles de bâtiment utilisés. Dans le second cas, deux solutions sont étudiées. La première consiste à ne considérer que la partie fiable des primitives extraites (ombres, textures). La reconstruction obtenue est correcte mais incomplète. En utilisant les échos forts, le bâtiment reconstruit est plus complet mais la reconstruction est bruitée.

6.4. Améliorations

Des perspectives pour améliorer la méthode de reconstruction sont proposées dans ce paragraphe. Ces améliorations concernent tout d’abord l’algorithme de génération d’hypothèses pour tenir compte des échos triples. Ensuite, une version itérative de la méthode de reconstruction est proposée pour valider une hypothèse, affiner la reconstruction et étendre la diversité des modèles de bâtiments.

6.4.1. Génération de contraintes sur les échos triples

Les échos triples, de faible amplitude, sont supposés n’être visibles que dans les ombres. Le cas où les ombres sont tronquées par des échos triples sol / mur / sol ou mur / sol / mur est étudié. Considérons une ombre délimitée par les segments FO (fin ombre, segment le plus éloigné du capteur, de hauteur h_{FO}) et DO (début ombre, segment le plus proche du capteur, de hauteur h_{DO}). Le type d’écho triple apparaissant dans les ombres peut être distingué par la hauteur du segment DO.

$$h_{DO} < 0$$

Il s’agit du cas où les échos triples sol / mur / sol, de hauteur négative, apparaissent dans l’ombre (cf. figure 153). Aucune hypothèse fiable ne peut être générée. Néanmoins, des contraintes sur la hauteur h et la longueur, selon l’axe des distances, w du bâtiment peuvent être déterminées.

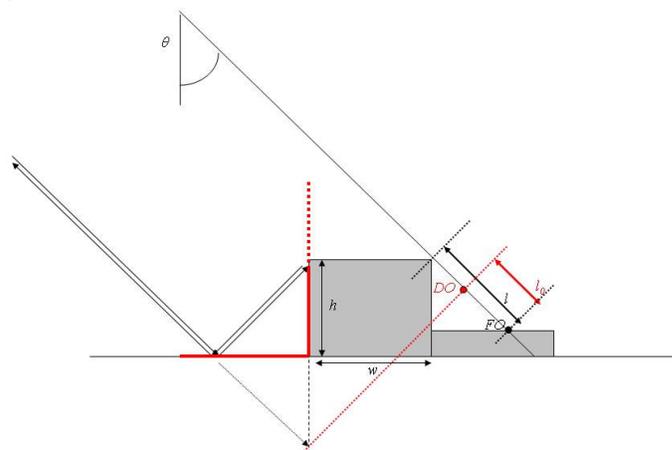


Figure 153 – exemple d’ombre tronquée par des échos triples sol / mur / sol

Les contraintes générées sur h sont obtenues par mesure monoscopique et par mesure interférométriques :

$$h \geq \max(l_0 \cos(\theta) + h_{FO}, -h_{DO}) \quad (103)$$

où θ est l'incidence de l'onde RADAR.

En utilisant la relation (63) p59, la contrainte sur w générée est :

$$w \leq \max(l_0 \cos(\theta) + h_{FO}, -h_{DO}) \frac{(2 \cos^2(\theta) + \sin(\theta) \cos^2(\beta))}{\cos(\theta) \sin(\theta)} \quad (104)$$

où θ et β sont respectivement l'angle d'incidence de l'onde et l'angle du bâtiment par rapport à l'axe des distances.

$h_{DO} = 0$

Ce cas concerne les ombres tronquées par des échos triples mur / sol / mur ou des échos doubles. La différenciation des deux cas est effectuée sur l'amplitude du segment DO . En effet, l'écho triple a une faible amplitude alors que l'écho double a une très forte amplitude. Le génération d'hypothèse lorsque un écho double tronque les ombres est présenté au §5.4.1, p149. Lorsque les échos triples mur / sol / mur tronquent l'ombre (cf. figure 154), ils ne fournissent aucune information sur la localisation et la hauteur de la façade.

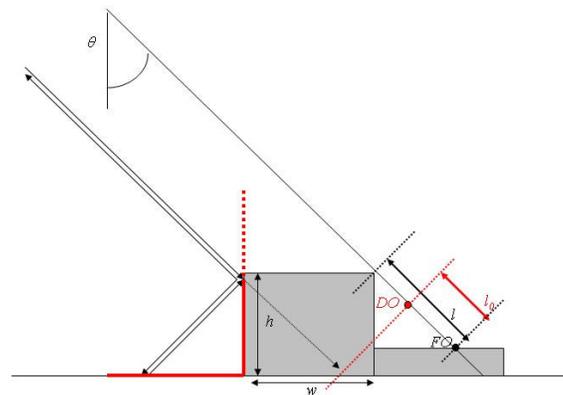


Figure 154 – exemple d'ombre tronquée par des échos triples mur / sol / mur

Dans ce cas, seule la mesure monoscopique est utilisée pour générer la contrainte :

$$h \geq l_0 \cos(\theta) + h_{FO} \quad (105)$$

6.4.2. Amélioration de la méthode de reconstruction

Pour améliorer les résultats de la reconstruction, un simulateur géométrique, permettant de projeter une scène de la géométrie terrain à la géométrie RADAR, peut être utilisé pour affiner le modèle [Bolter 2001], [Soergel 2003a].

Les primitives simulées peuvent être utilisées pour :

- valider une hypothèse
- améliorer l'estimation de la localisation planimétrique et de la hauteur
- étendre les modèles de bâtiments à des toits en pentes ou à double pentes

validation d'une hypothèse :

La simulation d'une hypothèse générée peut faire apparaître des primitives qui non pas été extraites. En effet, la figure 155a présente l'hypothèse générée pour l'exemple du bâtiment avec un patio. L'image issue de la simulation géométrique de cette hypothèse (cf. figure 155b) est composée de quatre couleurs :

- le gris foncé pour le sol
- le gris clair pour les façades avant et les toits
- le blanc pour les échos forts (le simulateur utilisé ne fournit que les échos délimitant les façades avant des bâtiments).
- le noir pour les ombres

La simulation permet de voir qu'une ombre devrait apparaître au niveau du patio. Après vérification cette ombre, entourée en rouge sur la figure 155c, est effectivement présente sur l'image d'amplitude. L'hypothèse générée serait donc validée. On pourrait alors imaginer relancer la méthode de reconstruction après avoir extrait cette ombre.

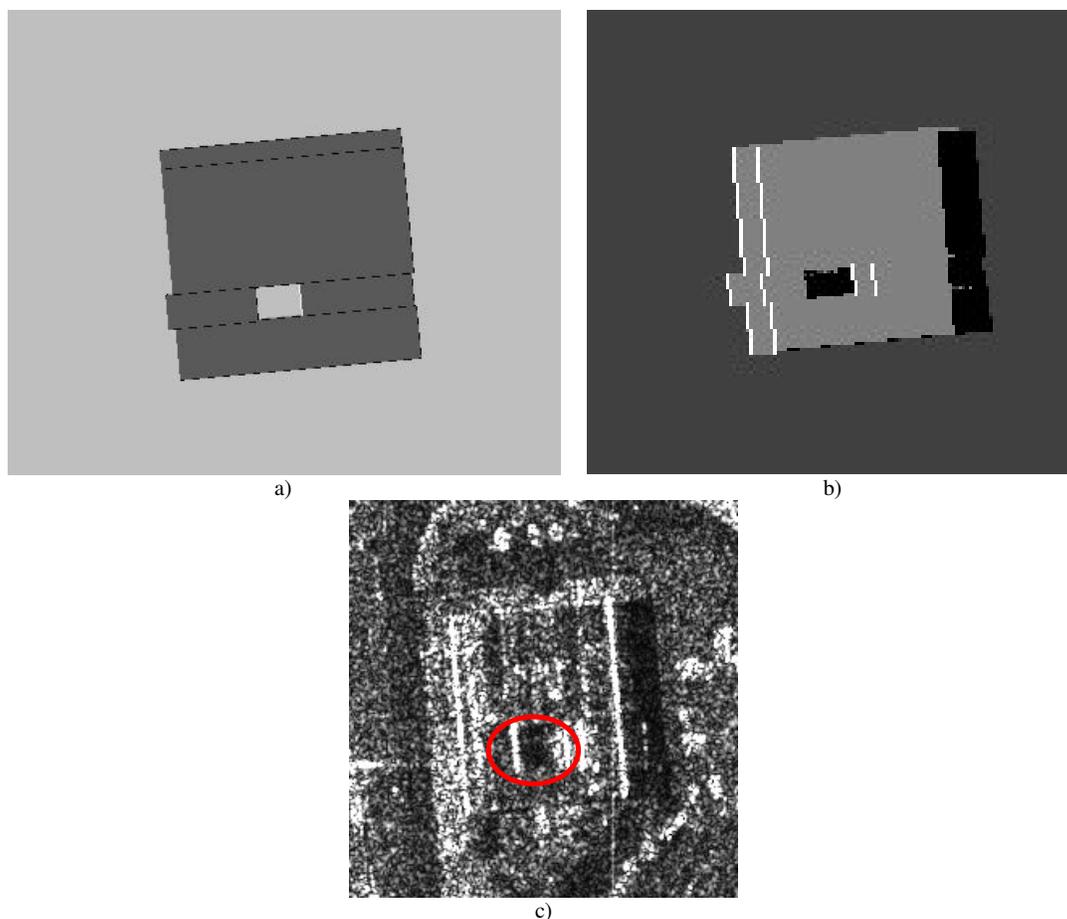


Figure 155 – a) hypothèse générée par la méthode de reconstruction,
b) simulation géométrique de l'hypothèse
(gris foncé pour le sol, gris clair pour les toits et les façades, blanc pour les échos, noir pour les ombres)
c) ombre sur la signature RSO permettant de valider l'hypothèse générée

affinement de la reconstruction :

Le simulateur géométrique peut également être utilisé pour affiner la reconstruction, en améliorant l'estimation de la localisation et de la hauteur du bâtiment. Ceci est possible en comparant le résultat de la simulation avec l'image RSO.

L'affinement est proposé pour l'interférométrie par détection et élimination des zones de repliement [Bolter 2001]. Pour cela, il serait nécessaire de faire évoluer le simulateur géométrique pour que les zones de repliement (c'est-à-dire les façades avant) apparaissent d'une autre couleur. De cette façon, seules les parties du toit où il n'y a pas de repliement pourraient être utilisées pour estimer la hauteur du toit.

Une analyse des différences entre l'image simulée et l'image RSO pourrait également être envisagée. En effet, en reprenant le même exemple que précédemment, la partie située juste avant le patio est estimée à une hauteur trop élevée (cf. §5.7, p167). La conséquence sur l'image simulée est de faire apparaître, sur cette partie, le bord du toit à une distance plus faible et de couper l'écho double et l'écho du bord du toit en plusieurs parties. Ce phénomène n'est pas visible sur l'image d'amplitude. Lors de l'étape d'affinement, une hauteur sur cette zone pourrait être déterminée de manière à ce que les échos de cette façade soient dans le prolongement de ceux des autres parties.

extension des modèles :

La reconstruction de bâtiments avec des modèles à toit en pente ou double pente est effectuée dans [Soergel 2003a] après avoir obtenu un modèle initial à toit plat. L'idée est d'utiliser ce modèle comme un *a-priori* fort pour estimer de façon plus précise la forme du toit.

Sur chaque zone homogène et pour chaque type de toit, le ou les plans, dont les paramètres minimisent l'erreur quadratique moyenne pourraient être calculées. Pour le modèle de toit en simple pente, un seul paramètre est nécessaire (la pente selon l'une des directions privilégiées). Pour le toit à double pente, deux paramètres sont nécessaires (pour les deux pentes). Pour chaque hypothèse issue de ces modèles, une simulation serait alors effectuée.

L'erreur quadratique moyenne sur la localisation entre les hypothèses simulées et les zones détectées serait ensuite calculée. Lors de ce calcul il ne faut pas utiliser les primitives extraites car l'étape d'extraction utilise les hypothèses de toit plat ou plat par morceaux. La comparaison des ombres extraites et des primitives simulées pour d'autres types de toit que les toits plats ou plats par morceaux n'est donc pas adaptée.

L'hypothèse minimisant la somme de ces deux erreurs serait retenue comme l'hypothèse la plus probable. Un affinement de la reconstruction comme proposé précédemment pourrait être effectué.

6.4.3. Conclusion

Des perspectives pour l'amélioration de la chaîne de reconstruction ont été proposées. Elles concernent tout d'abord l'étape de gestion d'hypothèses pour déterminer les contraintes associées aux échos triples. Ces contraintes supplémentaires pourraient aider à favoriser certains modèles. Ensuite, plusieurs pistes pour améliorer la méthode de reconstruction grâce à un simulateur géométrique sont envisagées. Elles pourraient permettre de valider / infirmer certaines hypothèses, affiner la reconstruction des bâtiments ou encore étendre le type de modèle.

6.5. Conclusion

Dans ce chapitre nous avons appliqué les contributions principales de cette thèse à d'autres jeux de données. Tout d'abord le modèle de mélange interférométrique a été validé sur données polarimétriques. Il a montré de bonnes performances sur les polarisations hH et vV qui permettent de caractériser les échos doubles et les réflexions simples. L'erreur quadratique moyenne sur l'estimation de hauteur est passée de 1.3m à 0.5m. Le modèle interférométrique a donc été validé sur un deuxième jeu d'images haute résolution.

Le modèle interférométrique a également été comparé à un modèle de mélange polarimétrique interférométrique. Ces deux modèles ont montré des performances similaires pour estimer la hauteur interférométrique des bâtiments à toit plat. De plus, dans ces deux modèles l'estimation de la hauteur des bâtiments est différente. En effet, dans le modèle interférométrique, la hauteur est calculée par rapport à une référence (écho double ou sol). Pour le modèle polarimétrique, la hauteur du bâtiment est dans la mesure de l'écart angulaire du champ de valeur de la cohérence. C'est pourquoi les résultats obtenus permettent également d'inter-valider les deux modèles.

Nous avons également montré que la méthode d'optimisation d'interférogramme n'est pas optimale sur les zones où il y a du repliement car le maximum de cohérence ne correspond pas forcément à une des contributions. C'est pourquoi cette technique ne doit être utilisée que sur les zones où il n'y a pas de repliement.

Puis la méthode de reconstruction a été utilisée sur des données interférométriques de meilleures résolutions. La conséquence de l'augmentation de la résolution est l'augmentation des primitives et donc du temps de calcul. Dans les deux exemples proposés, la méthode proposée a permis d'obtenir une reconstruction en accord avec les modèles de bâtiments choisis. Les résultats ont montré que la reconstruction est géométriquement correcte lorsque les primitives extraites sont suffisantes pour décrire les extrémités du bâtiment. Lorsque les primitives ne sont pas extraites correctement, la reconstruction n'est pas bonne.

Enfin, des perspectives pour améliorer la méthode de reconstruction ont été dégagées. Elles concernent tout d'abord la prise en compte des échos triples dans la méthode de génération d'hypothèse. Même si les échos triples n'apportent pas d'informations suffisantes pour générer des hypothèses, ils pourraient néanmoins être utiles pour la validation d'hypothèses. De plus, en nous inspirant des travaux existants, nous avons détaillé des algorithmes, basés sur l'utilisation d'un simulateur géométrique, qui permettraient d'obtenir une reconstruction plus précise et d'étendre la diversité des modèles de bâtiments.

Conclusion générale

Dans cette thèse une méthode de reconstruction de bâtiments sur images RSO interférométrique haute résolution en bande X a été proposée. Les modèles de bâtiments ont été choisis à empreinte à angles droits et à toit plat par morceaux car ils permettent de décrire un large spectre de bâtiments des villes ou zones industrielles.

Pour réaliser cet algorithme de reconstruction, plusieurs étapes ont été détaillées tout au long de cette thèse.

Tout d'abord une analyse des signatures de bâtiments sur images RSO interférométriques a été menée. Elle a permis de mettre en évidence les difficultés d'interprétation liées principalement à l'acquisition en distance et aux multiples réflexions de l'onde RADAR. Les échos doubles apparaissent à l'altitude du sol et permettent de délimiter la base du bâtiment. Les échos qui subissent plus de deux réflexions / rétrodiffusion apparaissent à des distances plus importantes que l'objet qui est physiquement à leur origine. Pour deux types d'échos triples, différenciés par la nature de la surface où a lieu la première réflexion, les localisations ainsi que les altitudes théoriques ont été validées.

En milieu urbain, à cause des nombreuses ruptures de relief, le repliement est très présent. Nous avons montré que dans les zones de repliement, l'information interférométrique correspond à un mélange d'altitude. Dans le cas d'un mélange de deux contributions, nous avons proposé un modèle de mélange interférométrique. Ce modèle a été validé sur deux jeux d'images sur lesquels il a permis d'améliorer l'estimation de hauteur interférométrique de bâtiments. Sur les images de Dunkerque et Avignon, le modèle interférométrique a permis de diminuer l'erreur quadratique moyenne sur l'estimation de hauteur interférométriques de bâtiments de respectivement 43% et 58%. Le modèle interférométrique a également été comparé à un modèle de mélange sur données polarimétriques interférométriques. Les deux modèles ont montré la même capacité à séparer les deux contributions du mélange.

Les précisions des différentes techniques pour extraire l'information 3D ont été également comparées. La mesure monoscopique sur les ombres est, au niveau de l'altimétrie, plus précise que la taille de la cellule de résolution. Néanmoins, à cause du bruit, du repliement et des réflexions triples, les ombres sont souvent bruitées et apportent parfois peu d'informations. La mesure monoscopique sur la parallaxe est légèrement moins précise. Elle nécessite d'avoir une information sur les deux extrémités de la façade. Cette mesure n'est également pas toujours utilisable car la signature des bâtiments ne présente pas toujours un écho franc ou des textures sur le bord du toit. Enfin, l'interférogramme présente la moins bonne précision altimétrique car les hauteurs d'ambiguïté sur les images sont de l'ordre de 200m. Son principal avantage est d'être une information dense. Dans les zones de repliement, lorsque plus de deux contributions sont mélangées, l'information interférométrique ne doit pas être prise en compte.

Une méthode de reconstruction ascendante de bâtiment a été ensuite proposée. A partir des images RSO, des traitements sont effectués pour extraire les signatures des bâtiments et pour obtenir, par gestion d'hypothèses, une ou plusieurs hypothèses conformes aux modèles de bâtiments à toits plats ou plats par morceaux.

Notre méthode de reconstruction consiste tout d'abord à filtrer les images RSO pour diminuer le chatouement. Pour cela le filtre mean shift a été introduit en imagerie RSO. Contrairement aux filtres paramétriques RSO usuels, il n'utilise aucun a priori sur les statistiques au sein des images RSO. Néanmoins, il a montré des performances similaires aux filtres paramétriques sur images synthétiques et sur images RAMSES. Une extension de ce filtre à l'image de cohérence, où les ombres sont également visibles a été proposée. Le mean shift étendu filtre alors simultanément l'image d'amplitude et de cohérence pour fournir, en sortie, deux images filtrées. Il a été ensuite montré que l'information, au sein des deux images filtrées, n'est pas strictement identique. Un algorithme de fusion a été proposé pour améliorer la détection des ombres.

A partir des images filtrées, des algorithmes d'extraction des primitives constituant les signatures des bâtiments, à savoir les échos forts, les textures et les ombres, ont été présentés ou proposés. Pour limiter le bruit sur les images, ces extracteurs utilisent les hypothèses effectuées sur les modèles de bâtiments. Une analyse géométrique a montré, pour les bâtiments conformes aux modèles choisis, que les échos et les contours des textures suivent deux directions privilégiées. Dans le cas des ombres, la direction de la visée doit être également prise en compte. C'est pourquoi l'algorithme d'extraction des ombres et textures est basé sur la transformée de Hough. Lorsque les primitives ne sont pas trop bruitées, il fournit de bons résultats. Les échos forts sont extraits, selon les deux directions privilégiées, par un algorithme de croissance de rectangle.

L'algorithme de reconstruction proposé est ensuite détaillé. Le bâtiment est tout d'abord séparé en zones homogènes (ou parties de bâtiment) où les mêmes primitives apparaissent. Sur chaque partie, les primitives sont ordonnées en géométrie terrain et les façades avant et arrière sont détectées lorsque les informations de hauteurs issues des mesures monoscopiques et interférométriques sont similaires. La reconstruction est effectuée, sur chaque zone homogène, par gestion d'hypothèses car les primitives extraites sont souvent insuffisantes pour décrire totalement la partie de bâtiment à reconstruire.

Toutes les hypothèses passant par les primitives extraites sont tout d'abord générées pour satisfaire localement les modèles de bâtiments choisis. Pour cela, une série de règles géométriques, différentes selon le type de primitives utilisées, a été proposée. Lorsque l'information n'est pas suffisante pour générer une hypothèse, des contraintes sont générées. Ensuite, l'étape de validation permet de sélectionner les hypothèses géométriquement valides et en accord avec les données interférométriques. Lorsque, sur des sections de zones homogènes, plusieurs estimations de hauteurs sont disponibles, elles sont fusionnées en tenant compte des écart-types sur chaque mesure. Les hypothèses de bâtiments complets proviennent de l'ensemble des combinaisons des hypothèses validées sur chaque zone homogène.

A l'issue de la validation, plusieurs hypothèses peuvent être obtenues. Une méthode de classement, basée sur la minimisation de la complexité stochastique, est proposée pour aider l'utilisateur à effectuer le choix final.

La méthode de reconstruction a été validée sur deux jeux d'images interférométriques RAMSES en bande X haute résolution. Les résolutions du premier jeu d'image sont de l'ordre de 50 cm. Le deuxième jeu possède des résolutions plus fines. Dans la majorité des exemples proposés, les algorithmes proposés ont permis d'obtenir une hypothèse géométriquement conforme à la vérité terrain. La précision planimétrique dépend de la qualité de l'extraction des primitives délimitant le bâtiment (échos doubles, échos du bord du toit et ombres). La précision altimétrique de la reconstruction dépend beaucoup du type de primitives extrait. Si une façade est détectée, la précision sur la mesure monoscopique permet d'obtenir une très bonne estimation de la hauteur du bâtiment. Dans certains cas, la hauteur du bâtiment est

estimée à 50cm près. Néanmoins, si l'information interférométrique est la seule disponible, l'erreur est plus importante. Il est important de noter que la méthode de reconstruction a permis d'obtenir des résultats intéressants sur les deux types d'images. Par conséquent elle se révèle peu sensible à l'amélioration de la résolution. Une utilisation de la méthode proposée sur des images très haute résolution pourrait donc être envisagée.

Enfin des perspectives d'amélioration de la méthode de reconstruction ont été proposées pour valider ou affiner les hypothèses reconstruites grâce à un simulateur géométrique. En effet, ce simulateur, permet de projeter les hypothèses de bâtiment en géométrie RADAR et de comparer le résultat de la reconstruction aux données. L'étude des similitudes ou différences entre signature simulées et signatures RADAR permet d'obtenir des indices sur la qualité du modèle reconstruit. Après avoir obtenu une ou plusieurs hypothèses de bâtiments à toit plat, une perspective est dégagée pour étendre le type de modèle aux bâtiments à toit en pente et double pente. Ainsi, un plus large spectre de bâtiments pourrait être reconstruit.

Dans cette thèse nous nous sommes intéressés uniquement à la reconstruction de bâtiments sur des imagerie. Pour représenter des scènes complètes en 3D, l'approche surfacique proposée dans [Tison 2004] a fourni des résultats intéressants, permettant d'obtenir une première ébauche de bâtiments sur une large zone. Ces deux approches sont donc complémentaires et on pourrait les utiliser conjointement. Les bâtiments obtenus sur le modèle numérique de surface pourraient servir de point de départ pour effectuer la reconstruction par la méthode proposée.

Il est intéressant de remarquer que la méthode de gestion d'hypothèses proposée pourrait également être adaptée à l'optique. En effet, sur ces images les textures et les ombres ont des origines différentes mais sont présentes. Les contours correspondants aux toits des bâtiments pourraient « remplacer » les échos forts. L'information 3D serait alors fournie par un couple stéréoscopique. En effet, en imagerie optique, l'appariement dense de scène est plus facile que sur images RSO car les images sont moins bruitées. Par conséquent une carte dense de disparité pourrait faire office d'interférogramme.

Publications personnelles

Cellier 2005, F. Cellier, H. Oriot, J.M. Nicolas, *Introduction of the mean shift algorithm in SAR imagery : application to shadow extraction*, EARSeL 2005, Porto, Juin 2005

Cellier 2006a, F. Cellier, H. Oriot, J.M Nicolas, *Study of altimetric mixture in layover areas on high resolution InSAR images*, EUSAR 2006, Dresde, Mai 2006

Cellier 2006b, F. Cellier, H. Oriot, J.M Nicolas, *Hypothesis management for building reconstruction from high resolution InSAR imagery*, IGARSS'06, Denver, Juillet 2006

Cellier 2006c, F. Cellier, E. Colin, *Building height estimation using fine analysis of altimetric mixtures in layover areas on polarimetric interferometric X-Band SAR images*, IGARSS'06, Denver, Juillet 2006

Bibliographie

Abdelfattah 1998, R. Abdelfattah, J.M. Nicolas, M.R. Boussema, *Topography optimization for accurate interferometric phase generation*, Proceedings of SPIE, SAR Image Analysis – Modeling and Techniques, Vol 3497, pp251-256, 1998.

Aizerman 1964, M. Aizerman, E. Braveman, L. Rozoncer, *Theoretical foundations of the potential function method in pattern recognition learning*, Automation & Remote Control Vol 25, pp821-837, 1964.

Baalz 2003, T. Balz, N. Haala, *SAR-based 3D reconstruction of complex urban environments*, ESA ESRIN Frascati Italy. January 2003.

Bamler 1993, R. Balmer, D. Just, *Phase statistics and decorrelation in SAR interferograms*, Proceedings of the International Geoscience and Remote Sensing Symposium, pp980-984, Tokyo, Japan, 1993.

Bolter 2001, R. Bolter, *Buildings from SAR : detection and reconstruction of buildings from multiple view high resolution interferometric SAR data*, PhD thesis from Technische Universität Graz, January 2001.

Boutry 1998, J.M. Boutry, *Image quality parameters of a high resolution experimental SAR*, ONERA –PUBLICATION- TP, Vol 204

Bors 2000, A.G. Bors, E.R. Hancock, R.C. Wilson, *Terrain Feature Identification by Modeling RADAR Image Statistics*, Proceedings IEEE International Conference on Image Processing, vol 1, pp721-724, Vancouver, Canada, 2000.

CCRS 2004, The Canada Center for Remote Sensing,
http://www.ccrs.nrcan.gc.ca/resource/tutor/stereo/chap5/chapter5_f.php

CCRS 2006, The Canada Center for Remote Sensing,
http://www.ccrs.nrcan.gc.ca/resource/tutor/polarim/chapter1.01_f.php

Cheng 1995, Y. Cheng, *Mean shift, mode seeking and clustering*, IEEE Transactions on pattern analysis and machine intelligence, Vol 17, n°8, August 1995.

Chien 1998, M. Chien, L. Yeh, Y. Yeh, F. Lin, *On geometrics properties of the numerical range*, Linear Algebra and its applications, 274, pp389-410, 1998

Clopinet 2006, I. Clopinet, www.clopinet.com/isabelle/Projects/SVM/applist.html

Cloude 1996, S. Cloude, E. Pottier, *A review of target decomposition theorems in radar polarimetry*, IEEE-GRS, vol. 34, n°2, pp 498-518, March 1996

Cloude 1998, S. Cloude, K. Papathanassiou, *Polarimetric SAR interferometry*, IEEE Trans. on Geoscience and Remote Sensing, vol 36, n°5, 1998.

Comanicui 2002, D. Comanicui, P. Meer, *Mean shift : a robust approach toward feature space analysis*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol24, n°5, May 2002.

Colin 2005, E. Colin, *Apport de la polarimétrie à l'interférométrie radar pour l'estimation des hauteurs des cibles et de paramètres de forêt*, Thèse de l'université Paris 6, 2005

- Dell'Acqua 2001*, F. Dell'Acqua, P. Gamba, U. Soergel, U. Thoenessen, *Fully 3D and joint 2D/3D segmentation of InSAR data over urban environments*, IEEE/ISPRS joint workshop on remote sensing and data fusion over urban areas, URBAN 2001. pp328-331.
- Desnos 1993*, Y.L. Desnos, V. Matteini, *Review on structure detection and speckle filtering on ERS-1 images*, EARSeL advances in remote sensing, Vol. 2, n°2-VI, 1993.
- Dreuillet 2006*, Dreuillet P., Cantalloube H., Colin E., Dubois-Fernandez P., Dupuis X. Fromage P., Garestier F., Heuze D., Oriot H., Peron J.L., Peyret J., Bonin G., du Plessis O.R. Nouvel J.F., Vaizan B., *The ONERA RAMSES SAR, latest significant results and future developments*, IEEE Conference on Radar 2006
- Eckstein 1995*, W. Eckstein, O. Munkelt, *Extracting objects from digital terrain models*, In T. Schenk, editor, Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes, Proc. Spie 2572, pages 222-231, 1995.
- Fjortoft 2003*, R. Fjortoft, Y. Delignon, W. Pieczynski, M. Sigelle, F. Tupin, *Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields*, IEEE Geoscience and Remote Sensing, Vol. 41, n°3, pp675-686, March 2003.
- Foerstner 1994*, W. Förstner, *A framework for low level feature extraction*, In Proc. Of the 3rd European Conference on Computer Vision, Sweden, 1994
- Frery 1999a*, A.C. Frery, C.da C. Freitas, S. J.S. Sant'Anna and C.D. Renno, *Statistical properties of SAR data and their consequences*, Seminars of the United Nations Programme on Space Applications. Vol.10 pp53-62, 1999.
- Frery 1999b*, A.C. Frery, C.da C.F. Yanasse, P.R. Vieira, S.J.S. Sant'Anna, C.D. Renno, *A user-friendly system for synthetic aperture RADAR image classification based on grayscale distributional properties and context*, Seminars of the United Nations programme on space application, vol 10, pp53-62.
- Fukanaga 1975*, K. Fukunaga, L. Hostetler, *The estimation of the gradient of a density function, with applications in pattern recognition*, IEEE Transactions on information theory, Vol 21, n°1, January 1975.
- Galland 2004*, F. Galland, *Partition d'images par minimisation de la complexité stochastique et grille active : application à la segmentation d'images de radar à ouverture synthétique*, Thèse de l'université d'Aix-Marseille III, décembre 2004.
- Gamba 1999*, P.Gamba, B. Houshmand, *Three dimensional urban characterization by IFSAR measurements*, Proc. of the 1999 IEEE International Geoscience and Remote Sensing Symposium (IGARSS'99), Vol. V, pp. 2401-2403, Hamburg, Germany, June 1999.
- Gamba 2000a*, P. Gamba, B. Houshmand, M. Saccani, *Detection and extraction of buildings from interferometric SAR data*, IEEE Transactions on Geoscience and Remote Sensing, Vol. 38, n. 1, pp. 611-618, Jan. 2000.
- Gamba 2000b*, P. Gamba, B. Houshmand, *Digital surface models and building extraction : a comparison of IFSAR and LIDAR data*, IEEE Transactions on Geoscience and Remote Sensing, vol.38, n°4, July 2000.
- Gattelli 1994*, F. Getelli, M. Guarnieri, F. Parizzi, P. Pasquali, C. Prati, F. Rocca, *The wavenumber shift in SAR interferometry*, IEEE Transactions on Geoscience and Remote Sensing, Vol 32, n°4, July 1994.

- Georgescu 2003, B. Georgescu, I. Shimshoni, P. Meer, *Mean shift based clustering in high dimensions : a texture classification example*, 9th International conference on computer vision , Nice, France, pp456-463, October 2003.
- Grenier 2005, T. Grenier, *Apport de l'espace des caractéristiques et des paramètres d'échelles adaptatifs pour le filtrage et la segmentation d'image*, Thèse de l'INSA Lyon, 2005.
- Henry 2003, C. Henry, *Détection de points brillants par corrélation complexe entre sous-vues d'une image RSO spatiale*, Institut national polytechnique de Toulouse, décembre 2003.
- Hoepfner 1999, K.B. Hoepfner, *Recovery of building structure from IFSAR-derived elevation maps*, CMPSCI Technical report 99-16, July 1999, Computer science department, Lederle graduate research center, University of Massachusetts, Amherst, MA 01003-4610.
- Jiang 1994, X. Jiang, H. Bunke, *Fast segmentation of range images into planar regions by scan line grouping*, Machine vision application, n°7, pp115-122, 1994.
- Just 1994, D. Just, R. Bamler, *Phase statistics of interferograms with application to synthetic aperture radar*, Applied optics, Vol.33, n°20, pp4361-4368, July 1994.
- Kirscht 1998, M. Kirscht, C. Rinke, *3D Reconstruction of buildings and vegetation from synthetic aperture radar (SAR) images*, Institut für theoretische nachrichtentechnik und informationsverarbeitung, Universität Hannover, Germany.
- Kohnle 1993, A. Kohnle, R. Neuwirth, W. Schuberth, K. Stein, D.H. Höhn, R. Gable, L. Hofmann, W. Euing, *Evaluation of essential design criteria for irst systems*, SPIE Vol 2020 Infrared technology XIX, 1993.
- Lopès 1990, A. Lopès, R. Touzi, E. Nezry, *Adaptative speckle filters and scene heterogeneity*, IEEE Transactions on Geoscience and Remote Sensing, Vol28, n°6, November 1990.
- Maître 2001, *Traitement des images de RSO*, sous la direction de H. Maître, Ed Hermes 2001.
- Michaelsen 2002, E. Michaelsen, U. Soergel, U. Stilla, *Grouping salient scatterers in InSAR data for recognition of industrial buildings*, IEEE 2002 Proceedings, Vol2 , pp613-616.
- Michaelsen 2005, E. Michaelsen, U. Soergel, U. Thoennesen, *Potential of building extraction from multi-aspect high resolution amplitude SAR data*, Joint workshop of ISPRS and the german association for pattern recognition, International Archives of Photogrammetry, Remote Sensing and the Spatial Information science, Vol 36, Part3/W24
- Nicolas 2003, J.M. Nicolas, *Une nouvelle application de la transformée de Mellin : l'étude des lois statistiques en imagerie cohérente. Applications aux images RSO*, Novembre 2003.
- Nicolas 2004, J.M. Nicolas, F. Tupin, *L'imagerie RADAR à synthèse d'ouverture*, Cours ENST Paris, Avril 2004
- Nocera 1996, L. Nocera, *Reconstruction du relief en imagerie RADAR à ouverture de synthèse : stéréoscopie et ombrages*, Université de Paris Jussieu 7, mai 1996
- Ortner 2002, M. Ortner , X. Descombes, J. Zerubia, *Building extraction from digital elevation model*, Rapport de recherche INRIA n°4517.
- Papathanassiou 1997, K. Papathanassiou, S. Cloude, *Polarimetric effects in repeat pass SAR interferometry*, IGARSS'97. Singapour, Aout 1997.
- Petit 2001, D. Petit, F. Adragna, J.D. Durou, *The filtering of layover area in high resolution IFSAR for the building extraction*, SAR Image Analysis, Modeling and Techniques III, Barcelone, Espagne, 2001

Quartulli 2001, M. Quartulli, M. Datcu, *Bayesian model based city reconstruction from high resolution ISAR data*. IEEE/ISPRS Joint workshop on remote sensing and data fusion over urban areas URBAN 2001. pp58-63.

Quartulli 2003a, M. Quartulli, M. Datcu, *Information extraction from high resolution SAR data for urban scene understanding*, 2nd GRSS/ISPRS Joint workshop on data fusion and remote sensing over urban areas. URBAN 2003, IEEE, pp115-119.

Quartulli 2003b, M. Quartulli, M. Datcu, *Information fusion for scene understanding from interferometric SAR data in urban environments*, IEEE Transactions on geoscience and remote sensing, Vol 41, n°9, September 2003.

Quartulli 2004, M. Quartulli, M. Datcu, *Stochastic geometrical modeling for built-up area understanding from a single SAR intensity image with meter resolution*. IEEE Transactions on geoscience and remote sensing, Vol 42, n°9, September 2004, pp1996-2003.

Radio 2006, <http://www.lpce.cnrs-orleans.fr>

Rissanen 1989, J. Rissanen, *Stochastic complexity in statistical inquiry*, World Scientific, Singapore, 1989

Silverman 1986, B.W. Silverman, *Density estimation for statistics and data analysis*, 1986 Ed. Chapman & Hall.

Simonetto 2002, E. Simonetto, *Extraction 3D de structures industrielles sur des images RAMSES haute-résolution par RADARgrammétrie*, thèse de l'université de Rennes 1, ONERA/ENST Bretagne, Avril 2002.

Soergel 2000, U. Soergel, U. Thoennesen, H. Gross, U. Stilla, *Segmentation of interferometric SAR data for building detection*, International archives of photogrammetry and remote sensing, Vol. 33, PartB1, pp328-335, Amsterdam 2000.

Soergel 2003a, U. Soergel, U. Thoennesen, U. Stilla, *Iterative building reconstruction from multi-aspect InSAR Data*, International Archives of Photogrammetry and Remote Sensing, Vol. 34, Part 3/W13 pp186-192, 2003.

Soergel 2003b, *Visibility analysis of man-made objects in SAR images*, 2nd GRSS/ISPRS Joint workshop on data fusion and remote sensing over urban areas. URBAN 2003, IEEE, pp120-124.

Soergel 2006, U. Soergel, E. Michaelsen, A. Thiele, U. Thoennesen, *Radargrammetric extraction of building features from high resolution multi-aspect SAR data*, IGARSS'06, Denver, 2006.

Stilla 2001, U. Stilla, U. Soergel, U. Thoennesen, *Potential and limits of InSAR data for the reconstruction of buildings*, IEEE/ISPRS joint workshop on remote sensing and data fusion over urban areas. URBAN 2001 pp64-68.

Stilla 2003a, U. Stilla, U. Soergel, U. Thoennesen, *Potential and limits of InSAR data for building reconstruction in built-up areas*, ISPRS Journal of Photogrammetry & Remote Sensing, pp113-123, January 2003.

Stilla 2003b, U. Stilla, E. Michaelsen, U. Soergel, K. Schulz, *Perceptual grouping of regular structures for automatic detection of man-made objects. Examples from IR and SAR.*, IGARSS'03.

Stilla 2005, U. Stilla, U. Soergel, U. Thoennesen, A. Brenner, *Potential and limits for reconstruction of buildings from high resolution SAR data of urban areas*, URSI GA 2005, New Delhi, India, 2005.

Steger 1996, C. Steger, *Extracting curvilinear structures :a differential geometric approach*, ECCV (1) pp630-641, 1996

Thiele 2006, A. Thiele, E. Cadario, K. Schulz, U. Thoennessen, U. Soergel, *Building recognition fusing multi aspect high resolution interferometric SAR data*, IGARSS'06, Denver, 2006.

Tison 2003, C. Tison, J.M. Nicolas, F. Tupin, *Accuracy of fisher distributions and log-moment estimation to describe amplitude distributions of high resolution SAR images over urban areas*, IGARSS'03, vol. III, pp 1999-2001, Toulouse, 21-25 July, 2003.

Tison 2004, C. Tison, *Interférométrie RSO à haute résolution en milieu urbain : application au calcul de MNS urbain*, PhD thesis from ENST, November 2003.

Tupin 1998, F. Tupin, H. Maître, J.F. Mangin, J.M. Nicolas, E.Pechersky, *Detection of linear features in SAR images : application to road network extraction*, IEEE Transactions on Geoscience and Remote Sensing, vol.36, n°2, pp434-453, March 1998.

Wand 1994, M.P. Wand, M.C. Jones, *Kernel smoothing*, Ed. Chapman & Hall/CRC, 1994.

Annexes

A. Présentation des données RAMSES

A.1. Présentation du capteur RAMSES

Les images utilisées dans le cadre de cette thèse sont issues du capteur RAMSES (RADAR Aéroporté Multi-Spectral d'Etude des Signatures) de l'ONERA sur un porteur de type Transall (cf. figure 156). Les antennes du capteur RAMSES sont situées sur la porte du Transall (entourée en rouge).



Figure 156 – Capteur RAMSES et son porteur : le Transall

Les caractéristiques du capteur sont rappelées dans le tableau ci-dessous ([Boutry 98], [Dreuillet 06]). La largeur de la bande utile, en bande X, est de 1200MHz. RAMSES peut donc acquérir des images RADAR haute et très haute résolution.

Bande	Fréquence centrale (GHz)	Largeur de Bande utile (Mhz)	Polarisation
P	0.43	75	Toutes
L	1.3	200	Toutes
S	3.2	300	Toutes
C	5.3	300	Toutes
X	9.5	1200	Toutes
Ku	14.3	1200	Toutes
Ka	35	1200	VV, HH
W	95	500	LR, LL

Tableau 1 - Capteur RAMSES et son porteur : le Transall

RAMSES utilise deux modes de fonctionnement pour échantillonner les données reçues par les antennes. La première est présentée au §1.2.5. Il s'agit du mode dérapé. Le signal

reçu est « déchirpé ». Ce mode permet d’obtenir, pour les fauchées couramment utilisée, un signal à échantillonner de bande inférieure à la bande utile. D’autres traitements sont ensuite nécessaires pour obtenir l’image. Le deuxième mode, plus classique, est appelé non-dérampé et consiste à numériser le signal reçu directement. Le convertisseur doit avoir une fréquence d’échantillonnage $F_{e,dist}$ supérieure à $2B_u$ où B_u est la bande utile. Le traitement consiste alors en l’utilisation du filtre adapté.

Le choix entre les deux types de numérisation est purement technique. En effet, il n’est pas toujours possible de trouver un convertisseur suffisamment rapide pour effectuer la numérisation sur toute la bande utile.

A.2. Présentation des images RSO interférométriques haute résolution RAMSES

Cette thèse s’inscrit dans le cadre de la reconstruction, les données interférométriques utilisées se présentent sous forme d’images de bâtiments isolés. Les bâtiments choisis sont de formes plus ou moins complexes. Un exemple de signature (image d’amplitude, interférogramme et image de cohérence) est présenté sur la figure 157.

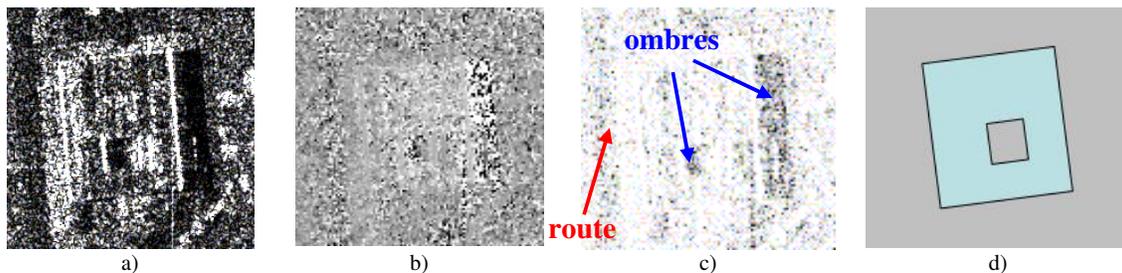


Figure 157 – a) Image en amplitude du couple interférométrique, b) interférogramme multivue (L=4), c) cohérence multivue (L=4) et d) représentation schématique du bâtiment

Sur l’image d’amplitude, figure 157a, la scène est difficilement interprétable pour tout non radariste. Ce bâtiment est de forme rectangulaire. Il est représenté sur le schéma de la figure 157d. Au centre, il y a un patio au niveau du sol. L’interférogramme se présente comme une carte dense de hauteurs et le toit du bâtiment apparaît plus clair que le sol, ce qui indique que cette partie de la scène est plus haute que le sol (figure 157b). En bande X et en acquisition monopasse, le capteur RAMSES fournit des images fortement cohérentes et seules les ombres et les zones peu rétrodiffusantes (routes) apparaissent avec une cohérence légèrement plus faible.

A.3. Estimation des paramètres des images

Estimation locale du rapport signal à bruit:

Le bruit sur les images RSO peut être estimé sur les zones où aucun signal n’est reçu. L’idéal est de choisir une zone où l’onde RADAR subit une réflexion miroir et n’est pas réfléchi vers le capteur. Ceci arrive sur une surface où il y a de l’eau stagnante par exemple. Les amplitudes moyennes, calculées par moyennage quadratique, sur cette zone et celle sur le clutter (sol) environnant, permettent d’estimer le rapport signal à bruit (RSB) local par :

$$RSB_{dB} = 20 \log_{10} \left(\frac{\bar{A}_{clutter}}{\bar{A}_{bruit}} \right) \quad (106)$$

Estimation de la hauteur d'ambiguïté

Pour calculer la hauteur d'ambiguïté il faut connaître les trajectoires des deux antennes, liées entre elles, du capteur RAMSES. Les deux images du couple interférométrique sont numérotées 1 pour l'image maître et 2 pour l'image esclave. Les coordonnées d'un pixel de l'image maître (en noir sur l'image 1) sont projetées dans l'image esclave pour deux hauteurs arbitraires (h_1 , en gris et h_2 , en rouge).

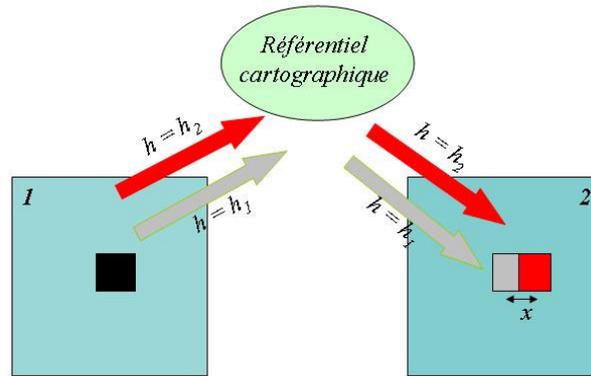


Figure 158 – Estimation de la hauteur d'ambiguïté

Soit x la différence de marche entre les deux projetés sur l'image esclave. La hauteur d'ambiguïté, notée h_{amb} , sous hypothèse d'onde plane, peut être obtenue par :

$$h_{amb} = \frac{\lambda(h_2 - h_1)}{x}$$

Si toutes les paramètres d'acquisition ne sont pas disponibles, une autre estimation de la hauteur d'ambiguïté peut être obtenue par (31), p27, en tenant compte du roulis pour déterminer l'angle d'incidence local.

Sur la figure 159, la hauteur d'ambiguïté estimée par (31) est tracée le long de la fauchée.

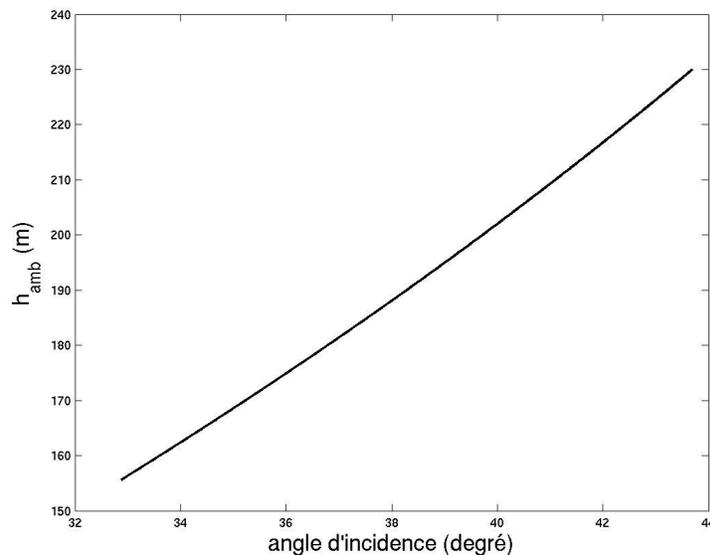


Figure 159 – Hauteur d'ambiguïté sur l'image de Dunkerque

Sur le site de Dunkerque, la base utilisée est faible et la hauteur d'ambiguïté au milieu de la fauchée est de l'ordre de 190m. L'avantage d'avoir une telle hauteur d'ambiguïté est qu'il n'y a pas besoin d'effectuer un déroulement de franges qui se révélerait compliqué en milieu urbain à cause des nombreuses discontinuités. Néanmoins, le bruit de phase sur les mesures interférométriques est très amplifié et sont peu précises (cf. §1.3.5).

B. Etude statistique des images en amplitude RAMSES sur Dunkerque

B.1. Introduction

De nombreuses études tentent de modéliser les distributions associées à la rétrodiffusion et à la réflexion de l'onde RADAR, et ce, par deux approches. La première se base sur la méthode de formation de l'image RSO. Les distributions employées ont alors une signification physique car elles se fondent sur la modélisation des phénomènes RADAR. La seconde est une approche empirique et consiste à chercher la distribution qui décrit le mieux et le plus simplement possible les images. La signification physique de ces lois n'est pas immédiate voire impossible. Ces modèles sont fréquemment utilisés par les algorithmes de filtrage, d'extraction ou de classification en imagerie RADAR. Dans un premier temps, les distributions des images RSO seront détaillées pour les zones homogènes. Puis dans le cas des zones hétérogènes, les modèles (physiques et empiriques) des lois communément utilisées seront rappelés.

B.2. Modèles statistiques issus de la physique RSO

B.2.1. Etude des lois statistiques des images RADAR sur une zone homogène

Ce paragraphe reprend les notions statistiques fondamentales exposées dans [Maître 2001] et décrit les distributions du chatoiement, de l'image complexe et de l'intensité dans une zone homogène.

B.2.1.1. Hypothèses du chatoiement pleinement développé

Contrairement aux techniques d'imagerie en éclairage incohérent (la lumière naturelle en imagerie optique), l'observation en imagerie RSO d'une zone homogène ne correspond pas à une image homogène. La nature cohérente de l'onde RADAR fait paraître l'image granuleuse, « bruitée ». Il s'agit du phénomène de chatoiement.

Pour une même cellule de résolution comportant N rétrodiffuseurs, le champ reçu par le capteur s'écrit comme la somme cohérente des ondes renvoyées par chacun des rétrodiffuseurs :

$$E = A e^{j\varphi} = \sum_{k=1}^N a_k e^{j\varphi_k}$$

où a_k et φ_k représentent l'amplitude et la phase de l'onde émise par le diffuseur k .

Sont alors établies les hypothèses du modèle gaussien circulaire complexe du chatoiement pleinement développé :

- le nombre de diffuseur N dans chaque cellule de résolution est suffisamment grand (> 5-6).

- l'amplitude et la phase du diffuseur k sont indépendantes de celles des autres diffuseurs.
- pour chaque diffuseur l'amplitude et la phase sont indépendantes entre elles.
- les variables a_k sont identiquement distribuées (suivant une loi normale ou uniforme par exemple).
- les phases ϕ_k sont uniformément distribuées entre $[-\pi ; \pi]$.

Pour N suffisamment grand et constant, d'après le théorème central limite, la partie réelle (en phase, indiquée \tilde{r}) et imaginaire (en quadrature, indiquée \tilde{i}) du signal reçu suivent une loi gaussienne de moyenne nulle et de même variance.

$$p_{\tilde{r}}(\tilde{r}|R) = \frac{1}{\sqrt{\pi R}} \exp\left(-\frac{\tilde{r}^2}{R}\right) \text{ et } p_{\tilde{i}}(\tilde{i}|R) = \frac{1}{\sqrt{\pi R}} \exp\left(-\frac{\tilde{i}^2}{R}\right)$$

avec $\sigma_{\tilde{r}, \tilde{i}}^2 = \sum_{k=1}^N \frac{1}{2} a_{\tilde{r}_k, \tilde{i}_k}^2 = \frac{1}{2} R$. R est appelé pouvoir réflecteur ou réflectivité de la cellule de résolution.

Les hypothèses faites sur N traduisent l'homogénéité de la zone. Elle se caractérise, dans chaque cellule de résolution, par le même type et le même nombre de rétrodiffuseurs. Par conséquent, une zone homogène peut se caractériser par une réflectivité constante. Si la réflectivité varie, la zone peut être considérée comme texturée.

B.2.1.2. Distribution des images en amplitude, intensité monovues

En posant $a = \sqrt{\tilde{r}^2 + \tilde{i}^2}$ on obtient la loi de Rayleigh qui est celle des images en amplitude monovue:

$$p_A(a|R) = \frac{2a}{R} \exp\left(-\frac{a^2}{R}\right)$$

avec $a \geq 0$, $E(a) = \sqrt{\frac{\pi R}{4}}$ et $E(a^2) = R$.

En posant $i = a^2$, on obtient la loi Gamma qui est celle des images en intensité monovue :

$$p_I(i|R) = \frac{1}{R} \exp\left(-\frac{i}{R}\right)$$

avec $i \geq 0$, $E(i) = R$ et $E(i^2) = R^2$.

B.2.1.3. Modèle multiplicatif du bruit

Pour les distributions des images monovues en amplitude et intensité, l'écart type de la variable est proportionnel à son espérance, d'où le nom de bruit multiplicatif associé au chatoiement noté S (Speckle en anglais). En effet, on a :

$$a = \sqrt{\frac{\pi R}{4}} S_a$$

avec $E(S_a) = 1$

et

$$i = R S_i$$

avec $E(S_i) = 1$.

Le chatoiement sur une image en amplitude suit le même type de loi que l'amplitude ou l'intensité mais avec une moyenne de 1. Ce modèle multiplicatif est valable uniquement dans le cadre du chatoiement complètement développé.

B.2.1.4. Distribution des images en amplitude, intensité multivues

Pour réduire l'aspect granuleux des images il est possible d'utiliser un traitement multivue. C'est à dire d'utiliser n vues non corrélées (sous-images) de l'image et d'effectuer un moyennage.

En posant $i_{ML} = \frac{1}{L} \sum_{k=1}^L i_k$ on obtient la loi Gamma généralisée suivante pour les images en intensité multivues:

$$p_{I_{ML}}(i_{ML}|R) = \left(\frac{L}{R}\right)^L \frac{1}{\Gamma(L)} \exp\left(-\frac{L i_{ML}}{R}\right) i_{ML}^{L-1}$$

avec $i_{ML} \geq 0$, $E(i_{ML}) = R$ et $E(i_{ML}^2) = \frac{L+1}{L} R^2$.

En posant $a_{ML} = \frac{1}{L} \sum_{k=1}^L a_k$, la distribution de l'amplitude n'a pas d'expression analytique.

On préfère alors utiliser le changement de variable $a_{ML} = \sqrt{i_{ML}} = \sqrt{\frac{1}{L} \sum_{k=1}^L a_k^2}$, on obtient alors la loi Rayleigh généralisée (ou loi de Nakagami) dont l'expression est :

$$p_{A_{ML}}(a_{ML}|R) = \left(\frac{L}{R}\right)^L \frac{2}{\Gamma(L)} a_{ML}^{2L-1} \exp\left(-\frac{L a_{ML}^2}{R}\right)$$

avec $a_{ML} \geq 0$, $E(a_{ML}) = \left(\frac{R}{L}\right)^{1/2} \frac{\Gamma(L + 1/2)}{\Gamma(L)}$ et $E(a^2) = R$.

B.2.2. Cas des zones hétérogènes

Dans le cas de zones hétérogènes, on ne considère plus la réflectivité constante mais comme une variable aléatoire. Cela apparaît dans les zones texturées lorsque le nombre de diffuseurs ne soit pas constant et/ou ces diffuseurs ne soient pas identiques.

La distribution Nakagami-Rice permet de caractériser les terrains tels que les grands reliefs [Bors 2000]. Ces terrains peuvent être considérés comme des zones hétérogènes. Les signaux en phase et quadrature sont modélisés comme gaussiens d'écart type σ . La partie imaginaire est considérée comme centrée et la partie réelle de moyenne μ : ceci revient à prendre une direction privilégiée pour la rétrodiffusion. Le même phénomène apparaît lorsqu'un réflecteur déterministe renvoie l'onde RADAR dans une direction privilégiée. La distribution de l'amplitude RADAR est donnée par :

$$p_A(a) = \frac{a}{\pi\sigma^2} \exp\left(-\frac{a^2 + \mu^2}{2\sigma^2}\right) I_0\left(\frac{a\mu}{\sigma^2}\right) \approx \frac{K}{\sigma} \sqrt{\frac{a}{2\pi\mu}} \exp\left(-\frac{(a-\mu)^2}{2\sigma^2}\right)$$

avec I_0 la fonction de Bessel modifiée de première espèce d'ordre 0
 K une constante qui vaut 1.06

Ce modèle est généralisé par Sijbers pour L pixels ou L vues aboutissant à la loi de Rice généralisée :

$$p_{A_L}(a) = \frac{a.L}{\sigma^2} \left(\frac{a}{a_0} \right)^{L-1} \exp\left(-L \frac{a^2 + a_0^2}{2\sigma^2}\right) I_{L-1}\left(\frac{a.a_0.L}{\sigma^2}\right)$$

En supposant que la texture suit une loi Gamma de paramètres μ et M et que le speckle suit une loi Gamme de paramètres 1 et L , l'intensité est distribuée suivant la loi K :

$$p_{I_L}(i) = \frac{1}{\Gamma(L)\Gamma(M)} \frac{2LM}{\mu} \left(\frac{LMu}{\mu} \right)^{\frac{M+L}{2}-1} K_{M-L} \left[2 \left(\frac{LMu}{\mu} \right)^{\frac{1}{2}} \right] \quad (107)$$

avec K fonction de Bessel de deuxième espèce.

B.3. Modèles empiriques

Les lois statistiques détaillées dans ce paragraphe n'ont pas de signification physique mais ont l'avantage de décrire correctement et simplement les distributions présentes en imagerie RSO.

La loi de Fisher (loi de Pearson de type VI) est utilisée pour décrire les statistiques des images RADAR [Nicolas 2003]. L'expression de cette loi pour les données en amplitude est la suivante :

$$p_A(a) = \frac{N}{M.\mu} \frac{\Gamma(N+M)}{\Gamma(N)\Gamma(M)} \frac{\left(\frac{N.a}{M.\mu} \right)^{N-1}}{\left(1 + \frac{N.a}{M.\mu} \right)^{N+M}} \quad (108)$$

Elle est présentée comme une extension de la loi Gamma (bien adaptée aux scènes naturelles). Ici, la texture est supposée suivre une loi Gamma Inverse, ce qui n'a aucune justification. Les paramètres N et M permettent d'avoir un comportement respectivement a tête lourde ('loi Gamma') et un comportement a queue lourde ('loi Gamma inverse') qui, en milieu urbain, permettent de bien décrire les zones d'échos forts en présence de sol [Tison 2003].

Dans [Simonetto 2002], les modèles empiriques suivants sont testés pour estimer la distribution des images en amplitude haute résolution en milieu urbain : Gauss, Lognormale, Weibull, Beta de première espèce et Beta de seconde espèce qui n'est autre que la loi de Fisher. La distance entre les distributions empiriques et estimées sont calculées par l'erreur quadratique moyenne (EQM) et le test de Kolmogorvo-Smirnov (TKS). L'auteur travaille en multivue ($L=4$) et conclut finalement que la loi Lognormale peut s'appliquer à toute l'image. Néanmoins, l'auteur estime les paramètres des lois par la méthode des moments ou du maximum de vraisemblance plutôt que la méthode des log-moments qui est plus adaptée pour l'estimation des lois à support positif [Nicolas 2003], [Tison 2003].

Dans [Frery 1999a] et [Frery 1999b], les auteurs se placent dans un contexte L vues et étudient la statistique d'images en intensité. La physique de l'imagerie RSO modélise le chatoiement par une loi Gamma. La texture est modélisée par la distribution inverse

généralisée d'une gaussienne. Le modèle multiplicatif donne la distribution en intensité $G_I(\alpha, \lambda, \gamma, L)$ de densité :

$$p_{i_{ML}}(i_{ML}) = \frac{L^L (\lambda/\gamma)^{\alpha/2}}{\Gamma(L) K_\alpha(2\sqrt{\lambda\gamma})} x^{L-1} \left(\frac{\gamma + Li_{ML}}{\lambda} \right)^{\frac{\alpha-L}{2}} \cdot K_{\alpha-n} \left(2\sqrt{\lambda(\gamma + Li_{ML})} \right) \quad , i_{ML} > 0 \quad , L > 0$$

et $\begin{cases} \gamma > 0, \lambda \geq 0 & \alpha < 0 \\ \gamma > 0, \lambda > 0 & \text{si } \alpha = 0 \\ \gamma \geq 0, \lambda > 0 & \alpha > 0 \end{cases}$

α est un facteur d'homogénéité. Plus la valeur absolue de ce paramètre est élevée, plus la zone observée est homogène. Ce modèle est général et deux cas particuliers sont étudiés :

- lorsque γ tend vers 0, la distribution converge en loi vers une loi $K_I(\alpha, \lambda, L)$.
- lorsque λ tend vers 0, la distribution converge en loi vers la loi $G_I^0(\alpha, \gamma, L)$ dont les propriétés sont décrites (utilisation de fonctions simples, elle peut être considérée comme le rapport de deux variables distribuées selon la loi Gamma, elle modélise les zones homogènes, hétérogènes et très hétérogènes, sa distribution cumulée est proportionnelle à la distribution de Snedecor).

B.4. Validation

Sur la figure suivante sont présentées des zones tests sur une première image. Nous avons choisi des zones d'ombre, de sol et d'échos forts relativement grandes et homogènes. Dans un premier temps, l'ombre est supposée ne pas contenir de réflexions multiples. Les zones utilisées pour l'estimation sont indiquées sur les images. Les mêmes ont été choisies pour effectuer l'estimation en multivue ($L=4$).

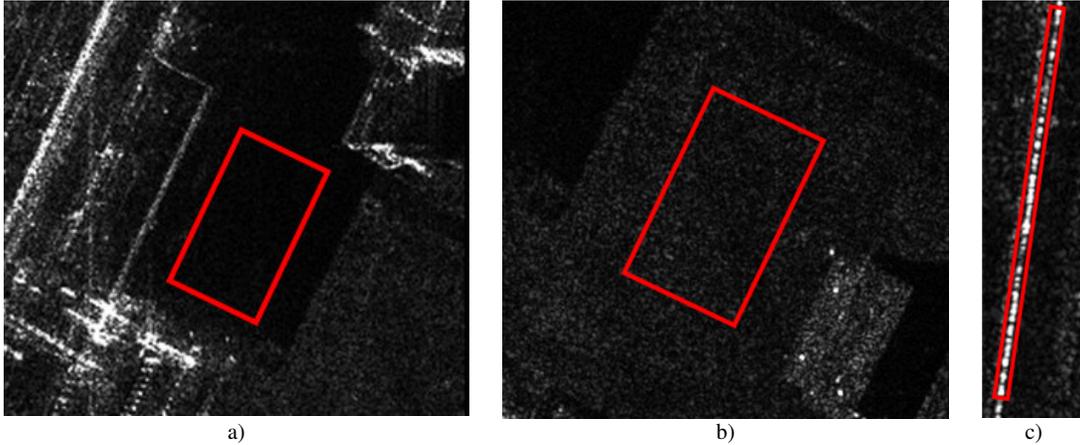


Figure 160 – Zones sur des images en amplitude correspondant à a) des ombres, b) du sol et c) un écho fort

Pour chaque zone, sont calculées les erreurs quadratiques moyennes (RMS) et le test de Kolmogorov-Smirnov (TKS) pour les estimations des statistiques sur les images monovues. Nous avons effectué, lorsque cela était possible, une estimation par plusieurs méthodes paramétriques usuelles : méthode des moments (MM), méthodes des log-moments (LogMM) et du maximum de vraisemblance (MV). Les parties supérieures et inférieures du tableau ci-après représentent les résultats d'estimation avec respectivement les lois empiriques et théoriques. En gras sont indiqués les modèles présentant les meilleurs résultats.

Zone	Ombre		Sol		Echos forts	
	Erreur RMS	TKS	Erreur RMS	TKS	Erreur RMS	TKS
Loi (Estimateur)						
Lognormale (MM)	8.47.10-3	0.060 (-)	0.89.10-3	0.060 (-)	2.08.10-3	0.112 (-)
Lognormale (MV)	7.44.10-3	0.062 (-)	0.83.10-3	0.067 (-)	1.48.10-3	0.043 (+)
Fisher (LogMM)	5.75.10-3	0.054 (-)	0.70.10-3	0.067 (-)	1.60.10-3	0.062 (-)
Weibull (LogMM)	3.79.10-3	0.013 (+)	0.36.10-3	0.009 (+)	2.12.10-3	0.091 (-)
Weibull Gén (LogMM)	6.59.10-3	0.053 (-)	0.74.10-3	0.057 (-)	1.60.10-3	0.042 (+)
Gamma Gén. (MM)	5.03.10-3	0.037 (-)	0.47.10-3	0.034 (-)	5.34.10-3	0.202 (-)
Gamma Gén. (MV)	3.97.10-3	0.018 (-)	0.37.10-3	0.012 (+)	2.80.10-3	0.165 (-)
Gamma Gén. (logMM)	3.85.10-3	0.015 (-)	0.36.10-3	0.011 (+)	2.46.10-3	0.119 (-)
K (LogMM)	4.73.10-3	0.032 (-)	0.56.10-3	0.046 (-)	1.73.10-3	0.070 (-)
Rayleigh Gén (LogMM)	3.85.10-3	0.015 (-)	0.36.10-3	0.011 (+)	2.46.10-3	0.119 (-)

Sur une autre zone de l'image les résultats en monovue sont :

Zone	Ombre		Sol		Echos forts	
	Erreur RMS	TKS	Erreur RMS	TKS	Erreur RMS	TKS
Loi (Estimateur)						
Lognormale (MM)	1.17.10-3	0.063 (-)	0.95.10-3	0.066 (-)	1.98.10-3	0.087 (-)
Lognormale (MV)	1.07.10-3	0.057 (-)	0.82.10-3	0.062 (-)	1.67.10-3	0.033 (+)
Fisher (LogMM)	0.83.10-3	0.052 (-)	0.61.10-3	0.052 (-)	1.63.10-3	0.041 (+)
Weibull (LogMM)	0.62.10-3	0.016 (+)	0.36.10-3	0.011 (+)	2.10.10-3	0.109 (-)
Weibull Gén. (LogMM)	0.96.10-3	0.048 (-)	0.72.10-3	0.052 (-)	1.67.10-3	0.042 (+)
Gamma Gén. (MM)	1.35.10-3	0.104 (-)	0.47.10-3	0.031 (-)	4.25.10-3	0.271 (-)
Gamma Gén. (MV)	0.66.10-3	0.026 (-)	0.37.10-3	0.014 (-)	2.76.10-3	0.182 (-)
Gamma Gén. (LogMM)	0.63.10-3	0.019 (+)	0.37.10-3	0.013 (+)	2.42.10-3	0.138 (+)
K (LogMM)	0.73.10-3	0.037 (-)	0.50.10-3	0.039 (-)	1.86.10-3	0.067 (+)
Rayleigh Gén. (LogMM)	0.63.10-3	0.019 (+)	0.37.10-3	0.013 (+)	2.42.10-3	0.138 (-)

L'estimation de statistiques telles que Rayleigh généralisée ou Fisher sur les zones d'ombre peut paraître à première vue surprenante. Néanmoins, il ne faut pas oublier que les ombres correspondent à une absence d'information reçue par le RADAR. Le capteur ne détecte alors que le bruit thermique. Ce bruit est supposé gaussien et sa quadrature l'est également. La loi Rayleigh généralisée en amplitude est donc théoriquement le modèle statistique des ombres pour les données en amplitude.

En monovue, aucune loi ne permet de décrire correctement les trois types de zones (ombre, sol et échos forts). Pour les ombres et le sol, les lois qui correspondent le mieux aux images RAMSES sont la loi Gamma généralisée sur les images en intensité (estimée par les logMM), Rayleigh généralisée sur les images en amplitude (estimée par le s logMM) et la loi de Weibull (estimée par les logMM). La correspondance de la loi Rayleigh généralisée (L est estimé proche de 1) et de nos données permet de vérifier que dans les zones homogènes le chatoiement est pleinement développé. La loi de Weibull n'a pas de justification physique pour cette bonne correspondance. Les résultats de l'estimation par la méthode des log-moments donne en revanche un paramètre de forme de 1.86 pour la zone d'ombre sur dk806. Or lorsque ce paramètre vaut 2, la loi de Weibull et la loi Rayleigh généralisée (avec $L = 1$) sont équivalentes. C'est pourquoi la loi de Weibull, grâce à son troisième paramètre s'adapte également mieux sur ces zones et présente toujours un TKS positif. On peut néanmoins s'étonner de voir la loi de Weibull avoir une erreur RMS plus faible que la loi de Rayleigh. Cela pourrait s'expliquer par la présence d'échos triples ou de lobes secondaires provenant d'échos du bâtiment. Par conséquent, il y a un mélange de lois au sein de la zone. Sur la zone sélectionnée, Même si la zone contient en grande majorité des pixels suivant la statistique des ombres, quelques pixels d'une autre statistique sont suffisant pour fausser les estimations. Cela pourrait également expliquer les faibles performances de la loi de Fisher. La présence d'échos triples alourdit la queue de distribution et fausse également l'estimation. Comme le modèle de Rayleigh est moins sensible à la queue de distribution, l'estimation de ses paramètres sur cette zone reste correcte.

L'estimation de la loi Gamma généralisée par le maximum de vraisemblance est meilleure que celle par la méthode des moments car le maximum de vraisemblance revient à utiliser les logmoments. En effet, pour estimer les paramètres de cette loi par la méthode des moments, il faut utiliser les moments d'ordre 1 et 2 en intensité et donc 2 et 4 en amplitude. Ce qui induit une variance de l'estimateur faisant intervenir les moments d'ordre 8 et donc une estimation moins fiable des paramètres de la loi. Par l'utilisation des log-moments, ceci n'apparaît plus car pour estimer les paramètres, le logarithme des données est utilisé et, par conséquent, seul un facteur multiplicatif apparaît ($\log I = 2 \cdot \log A$). Pour la même raison que précédemment, nous avons les mêmes résultats pour l'estimation de la loi Gamma généralisée (sur les intensités) et Rayleigh généralisée (sur les amplitudes) par la méthode des log-moments.

Sur les échos forts, la loi lognormale estimée par le maximum de vraisemblance et la loi de Weibull généralisée correspondent bien aux données. La loi de Fisher estimée par les log-moments pourrait également être envisagée. Il est difficile d'apporter une justification aux bonnes ou mauvaises estimations de nos données par ces lois car elles ont été choisies empiriquement. Ces tests ont été menés sur diverses zones des deux images et ont donné les mêmes résultats. Parfois certaines lois comme la loi K ou la loi Gamma estimée par le MV ont de meilleurs résultats mais les lois commentées précédemment sont les plus « stables ».

Nous nous intéressons maintenant à l'étude des lois statistiques sur les images RAMSES multivue pour les mêmes zones d'ombre, de sol et les échos forts.

Zone	Ombre		Sol		Echos forts	
	Erreur RMS	TKS	Erreur RMS	TKS	Erreur RMS	TKS
Loi (Estimateur)						
Lognormale (MM)	0.74.10-3	0.016 (+)	0.78.10-3	0.018 (+)	2.41.10-3	0.080 (+)
Lognormale (MV)	0.74.10-3	0.015 (+)	0.76.10-3	0.016 (+)	2.26.10-3	0.046 (+)
Fisher (LogMM)	0.72.10-3	0.008 (+)	0.76.10-3	0.010 (+)	2.24.10-3	0.045 (+)
Weibull (LogMM)	0.96.10-3	0.068 (-)	1.04.10-3	0.064 (-)	2.67.10-3	0.134 (-)
Weibull Gén. (LogMM)	0.72.10-3	0.010 (+)	0.76.10-3	0.009 (+)	2.31.10-3	0.063 (+)
Gamma Gén. (MM)	0.95.10-3	0.056 (-)	0.97.10-3	0.059 (-)	3.19.10-3	0.302 (-)
Gamma Gén. (MV)	0.84.10-3	0.058 (-)	0.89.10-3	0.044 (-)	2.93.10-3	0.225 (-)
Gamma Gén. (LogMM)	0.82.10-3	0.047 (-)	0.88.10-3	0.042 (-)	2.84.10-3	0.157 (-)
K (LogMM)	0.72.10-3	0.013 (+)	0.76.10-3	0.010 (+)	2.52.10-3	0.086 (-)
Rayleigh Gén. (LogMM)	0.82.10-3	0.047 (-)	0.88.10-3	0.042 (-)	2.84.10-3	0.157 (-)

Pour la deuxième image, les résultats sont les suivants :

Loi (Estimateur)	Ombre		Sol		Echos forts	
	Erreur RMS	TKS	Erreur RMS	TKS	Erreur RMS	TKS
Lognormale (MM)	1.19.10-3	0.020 (+)	0.69.10-3	0.019 (+)	2.93.10-3	0.103 (-)
Lognormale (MV)	1.18.10-3	0.020 (+)	0.70.10-3	0.017 (+)	2.91.10-3	0.085 (+)
Fisher (LogMM)	1.17.10-3	0.016 (+)	0.68.10-3	0.013 (+)	2.81.10-3	0.049 (+)
Weibull (LogMM)	1.49.10-3	0.083 (-)	0.91.10-3	0.064 (-)	3.29.10-3	0.171 (-)
Weibull Gén. (LogMM)	1.19.10-3	0.029 (+)	0.68.10-3	0.010 (+)	3.01.10-3	0.112 (-)
Gamma Gén. (MM)	1.71.10-3	0.095 (-)	0.87.10-3	0.057 (-)	3.71.10-3	0.304 (-)
Gamma Gén. (MV)	1.38.10-3	0.081 (-)	0.78.10-3	0.046 (-)	3.49.10-3	0.240 (-)
Gamma Gén. (LogMM)	1.32.10-3	0.065 (-)	0.76.10-3	0.039 (-)	3.41.10-3	0.190 (-)
K (LogMM)	1.19.10-3	0.031 (+)	0.68.10-3	0.009 (+)	3.09.10-3	0.116 (-)
Rayleigh Gén. (LogMM)	1.32.10-3	0.065 (-)	0.76.10-3	0.039 (-)	3.41.10-3	0.190 (-)

En multivue, la loi Gamma généralisée estimée par les log-moments n'est plus la mieux adaptée. Théoriquement, il est montré que pour des pixels non corrélés, les images multivues en amplitude suivent une loi Gamma lorsque le chatoiement est pleinement développé. Le nombre de vues estimé par la méthode des log-moments sur une zone d'ombre de l'image 1 est de 1.95. S'il n'y avait pas de corrélation entre les pixels, le nombre de vue devrait être de 4 (multivue 2x2). Ceci traduit donc la forte corrélation des images RAMSES.

La loi de Weibull est bien adaptée à nos données lorsque le nombre de vue est de 1. La loi de Weibull Généralisée, grâce à un paramètre supplémentaire, est bien adaptée aux données multivues. Il en résulte une estimation correcte des densités empiriques sur les zones d'ombre et de sol. L'estimation des échos forts s'avère néanmoins incertaine. On peut remarquer que, de la même façon que la loi de Weibull faisait le lien entre la loi Gamma et la loi Rayleigh, la loi de Weibull généralisée fait le lien entre la loi Gamma généralisée et Rayleigh généralisée dans le cas multivue.

La loi lognormale présente de meilleurs résultats que dans le cas monovue et son estimation par le maximum de vraisemblance décrit correctement les trois types de primitives sur les deux images. Ceci paraît normal car il a souvent été remarqué que, le nombre de vue augmentant, les lois sur les images RADAR « tendaient » vers une loi lognormale.

Sur les tests effectués, la loi de Fisher estimée par les log-moments correspond le mieux aux données et permet de décrire les zones d'ombre, de sol et d'échos forts ce qui rejoint [Tison 2003]. La loi de Fisher, avec ses trois paramètres, permet de modéliser le comportement à queue lourde et à tête lourde des lois en multivues.

Enfin, la loi K permet de décrire les zones d'ombres et de sol correctement mais pas les échos forts.

Ces tests ont également été effectués sur plusieurs zones appartenant à des images différentes. Il apparaît que pour les zones de sol, la loi de Weibull Généralisée ou la loi K correspondent parfois mieux que la loi de Fisher. En revanche, dans les autres zones, la loi de Fisher reste la plus adaptée.

C. Généralités sur les champs de Markov

Les champs de Markov sont très utilisés en classification et restauration d'image. L'image est alors associée à une réalisation d'un champ aléatoire. Tout comme les filtres, les champs de Markov tiennent compte du voisinage mais utilisent également l'information de connexité. Une hypothèse des champs de Markov est que la valeur de chaque pixel dépend uniquement de la valeur des pixels voisins. Par le théorème de Hammersley-Clifford, il est donc possible de relier ce champ de pixel à un champ de potentiel de Gibbs. Les potentiels sont alors reliés à des probabilités.

Pour trouver la configuration optimale, un recuit simulé est associé à un échantillonneur (Gibbs, Métropolis). A chaque itération, les échantillonneurs de Gibbs et Métropolis fournissent un tirage aléatoire d'un pixel selon une loi uniforme pour Métropolis et une loi définie par le voisinage pour Gibbs, et le pixel se voit attribuer une nouvelle classe. Le calcul de l'énergie associée à cette nouvelle configuration permet de savoir si elle est plus probable (énergie plus faible) que la précédente. Si ce n'est pas le cas, une autre classe pour le pixel est tirée dépendant des probabilités conditionnelles de ses voisins. Au début de la classification, la température associée au recuit simulé est élevée et les modifications importantes de configurations sont possibles. Puis, à chaque itération, l'algorithme baisse la température ce qui autorise des changements de plus en plus restreints. La condition d'arrêt porte sur le nombre d'itérations ou le nombre de pixels modifiés entre chaque itération. L'algorithme ICM (Iterative Conditional Mode) teste toutes les configurations possibles en chaque site de l'image et garde celle qui minimise l'énergie. Il faut préciser que cette méthode est itérative et chaque site de l'image est parcouru plusieurs fois. La convergence s'effectue dans ce cas vers un minimum local d'énergie et il faut donc s'assurer d'avoir un état initial proche du minimum global.

Les chaînes de Markov, contrairement aux champs, travaillent sur un voisinage unidimensionnel. L'application des chaînes aux données 2D en essayant de limiter la perte d'information peut se faire par l'utilisation du parcours de Peano. Il s'agit d'une façon (parmi d'autres) de représenter le voisinage spatial de l'image 2D en une seule dimension. Cette technique a été utilisée dans [Simonetto 2002] avec l'algorithme ICE-MPM-GEMI (Iterative Conditional Estimation – Maximum Posterior Mode – Generalized Mixture). La classification est faite sur trois classes (échos forts, sol, pixels sombres). Cet algorithme permet de tenir compte non pas d'une statistique mais d'un ensemble de statistique (GEMI). La statistique choisie pour chaque classe est celle qui minimise l'erreur quadratique moyenne entre l'histogramme empirique et théorique des pixels de la classe. L'affectation d'une classe à chaque pixel est obtenue par la configuration qui maximise la probabilité marginale à posteriori (MPM). L'approche ICE permet de mettre à jour les matrices de transitions entre chaque classe à l'issue de chaque itération.

Une comparaison des champs et chaînes de Markov sur des images JERS basse résolution peut être trouvée dans [Fjørtoft 2003]. Les résultats paraissent plus bruités sur la classification par chaîne de Markov car le voisinage est moins bien pris en compte. Néanmoins, le temps de calcul pour les chaînes de Markov est beaucoup moins important.

D. Résultats complémentaires sur le mean shift en imagerie RSO

D.1. Performances du mean shift sur image RADAR synthétiques

Les probabilités de détection des filtres médian, mean shift, Lee, Kuan, Frost amélioré, Gamma MAP et Fisher MAP sont consignées dans les tableaux ci-dessous pour des probabilités de fausse alarme de 1%, 2% et 5% sur images RADAR synthétiques monovues et multivues.

sans filtrage :

p_d en % pour différentes pfa sur les images synthétiques non filtrées

	Monovue				Multivue 2x2			
Contraste	3 dB	1.76 dB	0.79 dB	0.41 dB	3 dB	1.76 dB	0.79 dB	0.41 dB
Pfa 1%	0.99	0.75	0.59	0.55	4.97	1.95	0.95	0.72
Pfa 2%	2.94	2.24	1.78	1.64	11.96	5.29	2.75	2.06
Pfa 5%	8.80	6.66	5.35	4.94	27.17	13.92	7.69	5.98

Sans filtrage, les probabilités de détection sur images monovues et multivues augmentent en fonction du contraste mais reste faibles. Pour 3dB de contraste en multivue, la probabilité de détection est de 27% seulement lorsque la probabilité de fausse alarme est de 5%.

filtres non paramétriques :

p_d en % pour différentes pfa sur les images synthétiques filtrées médian 3x3

	Monovue				Multivue 2x2			
Contraste	3 dB	1.76 dB	0.79 dB	0.41 dB	3 dB	1.76 dB	0.79 dB	0.41 dB
Pfa 1%	6.60	2.40	1.049	0.73	64.70	17.97	3.16	1.38
Pfa 2%	15.33	6.30	2.99	2.15	81.66	32.65	7.81	3.75
Pfa 5%	32.72	15.94	8.30	6.22	93.03	53.63	17.86	9.84

p_d en % pour différentes pfa sur les images synthétiques filtrées mean shift dans le domaine spatial/intensité avec $h_A=2$ et $h_S=5$

	Monovue				Multivue 2x2			
Contraste	3 dB	1.76 dB	0.79 dB	0.41 dB	3 dB	1.76 dB	0.79 dB	0.41 dB
Pfa 1%	55.16	19.94	4.13	1.63	99.86	98.97	59.51	14.83
Pfa 2%	78.53	36.98	9.74	4.23	99.98	99.67	74.82	26.11
Pfa 5%	93.74	60.60	21.56	11.13	100	99.93	88.20	43.96

Le filtrage mean shift et, dans une moindre mesure, le filtre médian 3x3, permettent d'améliorer la détection. Pour une probabilité de fausse alarme de 2%, en multivue, les

probabilités de détection sont de 99.98% et 78.53% sur les images filtrées respectivement par le mean shift et le médian 3x3.

filtres paramétriques :

p_d en % pour différentes pfa sur les images synthétiques filtrées par Lee

Contraste	Monovue				Multivue 2x2			
	3 dB	1.76 dB	0.79 dB	0.41 dB	3 dB	1.76 dB	0.79 dB	0.41 dB
Pfa 1%	92.74	29.78	3.29	1.28	99.75	98.72	59.37	13.71
Pfa 2%	98.23	62.20	10.82	4.12	99.97	99.57	75.65	25.75
Pfa 5%	99.41	87.16	29.47	12.69	99.99	99.91	88.96	44.50

p_d en % pour différentes pfa sur les images synthétiques filtrées par Kuan

Contraste	Monovue				Multivue 2x2			
	3 dB	1.76 dB	0.79 dB	0.41 dB	3 dB	1.76 dB	0.79 dB	0.41 dB
Pfa 1%	74.71	14.51	2.12	1.02	98.24	74.99	5.92	1.78
Pfa 2%	92.33	41.27	7.41	3.36	99.47	92.60	22.25	5.92
Pfa 5%	97.64	72.60	22.16	10.60	99.86	97.72	55.60	19.82

p_d en % pour différentes pfa sur les images synthétiques filtrées par Frost

Contraste	Monovue				Multivue 2x2			
	3 dB	1.76 dB	0.79 dB	0.41 dB	3 dB	1.76 dB	0.79 dB	0.41 dB
Pfa 1%	99.22	72.21	12.32	3.28	99.69	98.75	63.53	15.77
Pfa 2%	99.67	86.09	23.56	7.62	99.97	99.59	78.12	27.80
Pfa 5%	99.88	94.85	41.45	17.42	100	99.93	90.06	46.18

p_d en % pour différentes pfa sur les images synthétiques filtrées par Gamma MAP

Contraste	Monovue				Multivue 2x2			
	3 dB	1.76 dB	0.79 dB	0.41 dB	3 dB	1.76 dB	0.79 dB	0.41 dB
Pfa 1%	94.25	37.81	4.27	1.51	98.57	95.13	49.97	10.65
Pfa 2%	97.68	66.03	12.29	4.53	99.12	97.03	68.11	22.60
Pfa 5%	98.86	87.03	30.32	13.07	99.39	98.19	81.77	40.85

p_d en % pour différentes pfa sur les images synthétiques filtrées par Fisher MAP

Contraste	Monovue				Multivue 2x2			
	3 dB	1.76 dB	0.79 dB	0.41 dB	3 dB	1.76 dB	0.79 dB	0.41 dB
Pfa 1%	31.99	11.39	3.60	1.99	96.13	37.37	5.97	2.84
Pfa 2%	42.47	16.91	6.13	3.66	97.83	62.36	10.65	5.24
Pfa 5%	63.45	28.30	12.38	8.26	99.06	90.50	23.50	12.06

Les filtres paramétriques permettent également d'améliorer la détection. On peut remarquer que, sur images monovues, les filtres de Frost amélioré et Gamma MAP présentent de meilleures performances que le mean shift. En effet, en monovue, le chatouillage est trop important et les statistiques sont difficiles à lisser. L'utilisation du multivue 2x2 permet de

diminuer l'aspect granuleux des images et d'améliorer les performances du mean shift. Les performances du filtre mean shift sont alors similaires voir meilleures que celles des filtres paramétriques.

D.2. Limitations du filtre médian 9x9 pour le filtrage des images RADAR synthétiques

Pour mettre en évidence les limitations du filtre médian sur une fenêtre 9x9, il a été comparé au filtre mean shift proposé par [Commaniciu 2002] sur l'image synthétique de la figure 161. L'image est divisée en deux zones de contraste 14dB. Chaque zone contient des blocs de 4x4 pixels de l'autre zone.

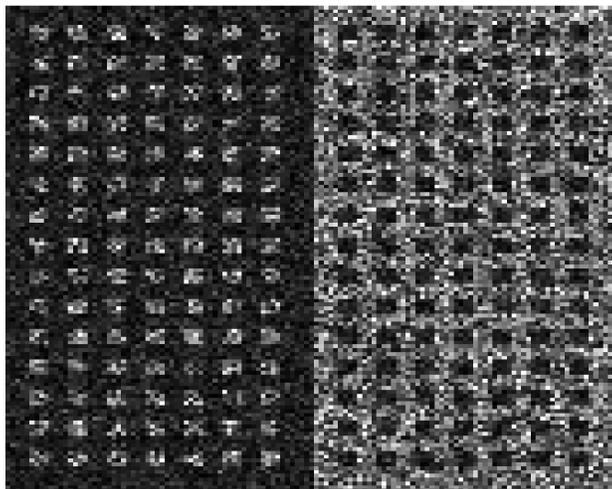


Figure 161 – Image synthétique « damier »

Les résultats du filtrage de cette image par le filtre médian 9x9 et le filtre mean shift sont représentés sur la figure 162. Sur l'image filtrée par le médian, tous les blocs de 4x4 pixels ont été supprimés. Sur images RSO, utiliser ce filtre signifierait perdre l'information associée aux points brillants. En revanche, l'image filtrée par le mean shift conserve une grande partie des blocs de 4x4 pixels. Certains sont néanmoins presque complètement supprimés et il n'en reste que les points les plus brillants.

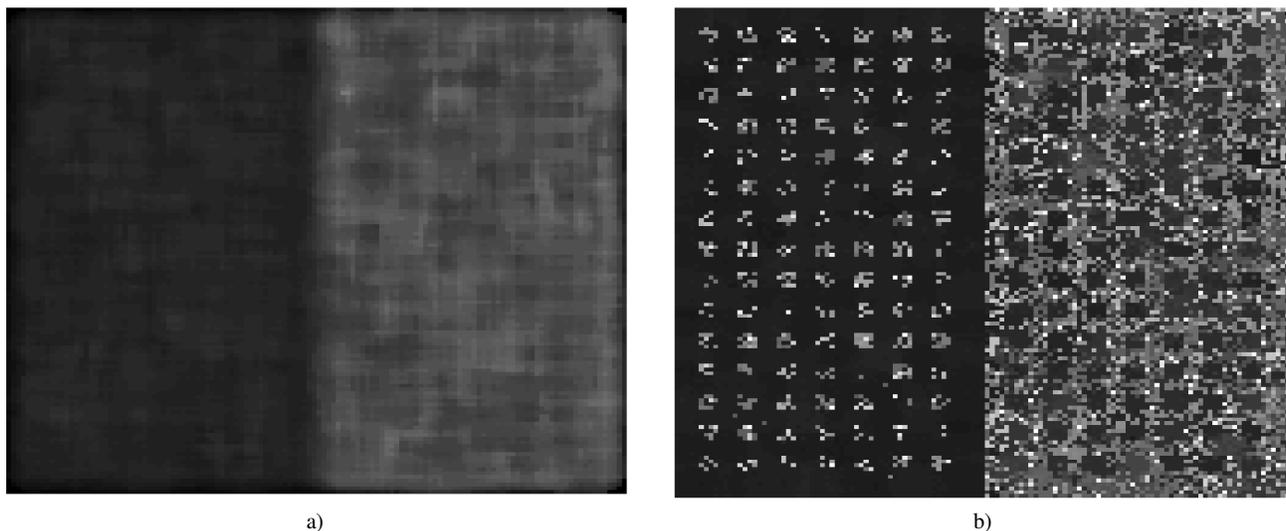


Figure 162 – filtrage de l'image synthétique « damier » par le filtre a) médian 9x9 et b) mean shift

Pour comparer quantitativement les filtrages, les courbes COR correspondant aux deux filtres sont tracées sur la figure 163. La courbe COR correspondant au mean shift est située bien au-dessus de celle du médian ce qui correspond à un meilleur filtrage.

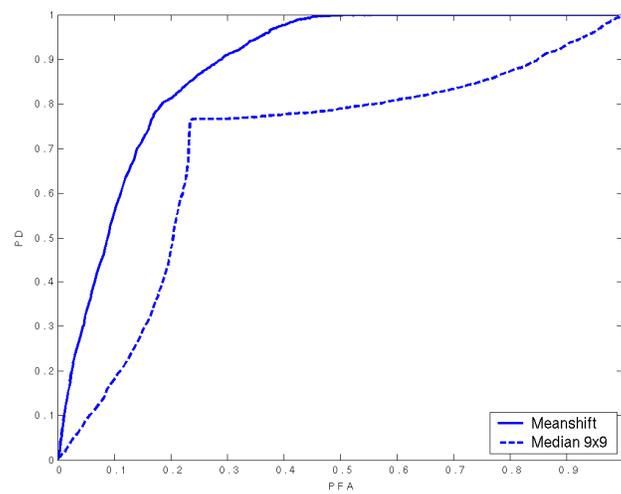


Figure 163 – Courbes COR des filtres mean shift et médian 9x9 sur l'image synthétique « damier »

E. Régularisation par programmation dynamique

E.1. Principe

Considérons un ensemble de donnée 1D, représenté en noir sur la figure 164, obtenu par un capteur. A cause du bruit sur la mesure et/ou d'éventuelles erreurs d'extraction, ces données ne correspondent pas à la réalité représentée en rouge. Le but de la régularisation est de débruiter les données extraites en tenant compte du voisinage spatial.

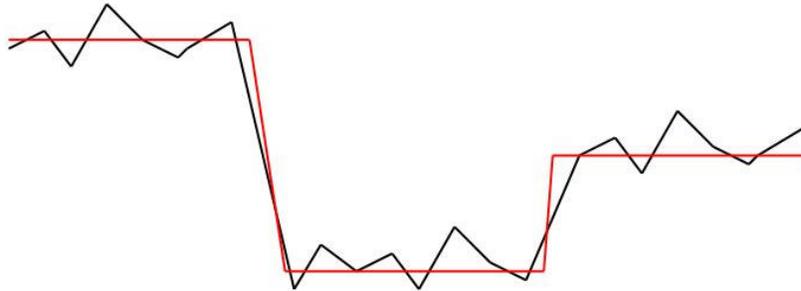


Figure 164 – Données bruitées (en noir) et réalité (en rouge)

Pour effectuer la régularisation, une fonction d'énergie, divisée en deux termes, est construite. Le premier terme est un terme d'attache aux données et permet d'assurer que les données régularisées soient proches des données initiales. Le second terme est le terme de régularisation et limite les trop fortes variations des données dans un voisinage spatial.

Soient $x_i, i=1\dots n$, les données bruitées et $y_i, i=1\dots n$ les données régularisées. Ces données sont supposées discrètes et appartenir à un ensemble Ω composé de m éléments. La fonction d'énergie est calculée par :

$$E = \Lambda f(x_0 - y_0) + \sum_{i=1}^n [\Lambda f(x_i - y_i) + g(y_i - y_{i-1})]$$

où f et g sont les fonctions d'énergie (généralement quadratiques) d'attache aux données et de régularisation.

Λ est un paramètre qui permet de faire varier la force de la régularisation.

La régularisation consiste alors à trouver la configuration des $y_i, i=1\dots n$, qui minimisent E .

E.2. Principe de la programmation dynamique

Le nombre total de configurations vaut n^m et peut être très important. Résoudre ce problème de minimisation en calculant l'énergie de toutes les configurations nécessiterait des temps de calcul trop importants. La solution de ce problème peut alors être obtenue en un

temps plus raisonnable par programmation dynamique. Pour cela, le problème de la régularisation est représenté sous forme d'un graphe en 2 dimensions où :

- $S_{i,j}$ représente un nœud de ce graphe correspondant à l'échantillon i et à l'élément j de Ω .
- $A_{j,j'}$ correspond à l'arc entre deux nœuds d'échantillons consécutifs reliant les éléments j et j' .

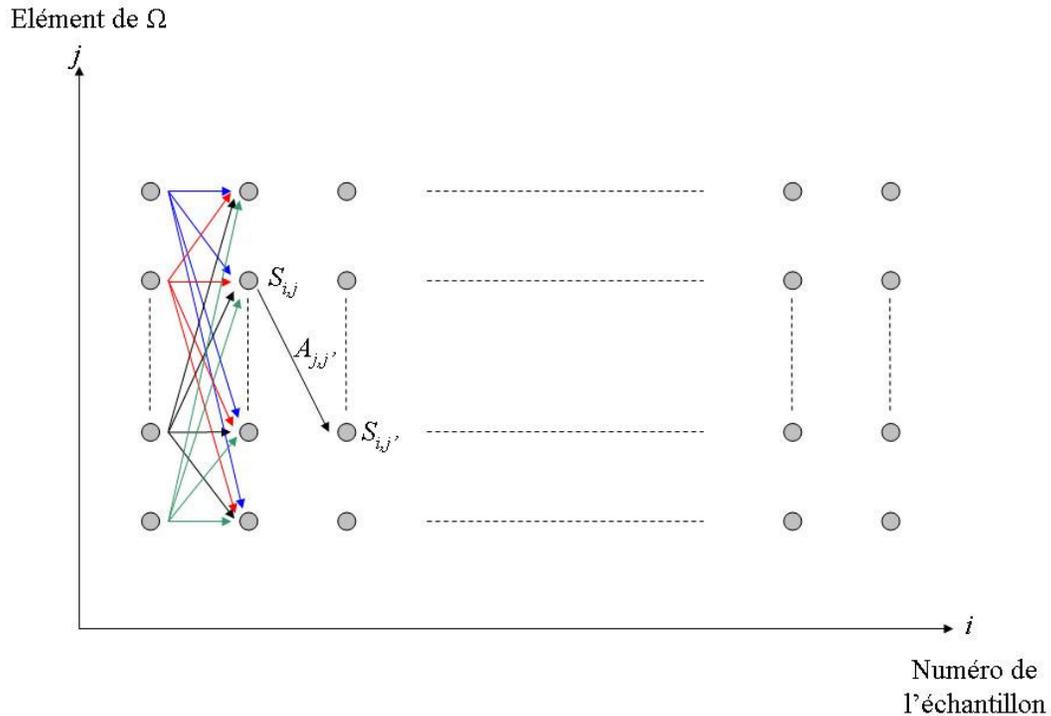


Figure 165 – Représentation du problème de régularisation sous forme d'un graphe

Une énergie est associée aux nœuds et arcs de ce graphe par :

- $E(S_{i,j}) = \Lambda f(x_i - y_j)$
- $E(A_{j,j'}) = g(y_j - y_{j'})$

L'algorithme de programmation dynamique est présenté ci-dessous :

Initialisation :

Calcul des énergies $E(S_{0,j}) = E_{0,j}$

Pour $i=1$ à n

Pour $j=1$ à m

Calculer l'énergie $E_{i,j} = E(S_{i,j}) + \min_{k=1..m} (E_{i-1,k} + E(A_{k,j}))$

Suivant j

Suivant i

Le chemin conservé est celui qui a l'énergie $\min_{j=1..m} (E_{n,j})$.

